University of Venda

# FORECASTING HOURLY SOLAR IRRADIANCE IN SOUTH AFRICA USING MACHINE LEARNING MODELS

By

**Mutavhatsindi Tendani**
**14013872**

submitted in fulfilment of the requirements for the degree of Master of Science in Statistics

in the

Department of Statistics
School of Mathematical and Natural Sciences
University of Venda
Thohoyandou, Limpopo
South Africa

Supervisor : Dr. Caston Sigauke

Co-Supervisor : Mr. Rendani Mbuvha

Date submitted: 12 August 2020

# Abstract

Solar irradiance forecasting is essential in renewable energy grids amongst others for back-up programming, operational planning, and short-term power purchases. This research study focuses on forecasting hourly solar irradiance using data obtained from the Southern African Universities Radiometric Network at the University of Pretoria radiometric station. This research project compares the predictive performance of long short-term memory (LSTM) networks, support vector regression (SVR), and feed forward neural networks (FFNN) model for forecasting short-term solar irradiance. While all the models outperform principal component regression model, a benchmark model in this study, the FFNN yields the lowest mean absolute error (MAE) and root mean square error (RMSE) on the testing set. According to findings, among the three fitted machine learning models, the FFNN model produced the best forecast accuracy based on MAE and RMSE. Forecast combination of machine learning models' forecasts is done using convex combination and quantile regression averaging (QRA). Based on MAE and average pinball losses, the QRA forecast combination model is the best forecast combination method, and also the best forecasting model compared with the individual machine learning models. Further analysis of the prediction interval widths (PIWs) based on the prediction interval coverage probabilities (PICPs), and prediction interval normalised average widths (PINAWs) including a count of the number of predictions below and above the PIs show inconsistency. The best model based on PINAW at 90% is QRA, FFNN at 95%, and SVR at 99%. There is inconsistency in the results of PICPs and PINAWs for different PINC values. The residual analysis shows FFNN as the best model with narrowest error distribution compared to other models, followed by QRA.

**Keywords**: *Forecast combination, machine learning, neural networks, solar irradiance forecasting, support vector regression.*

# Declaration

I, Tendani Mutavhatsindi, [student Number: 14013872], hereby declare that the dissertation titled: “Forecasting Hourly Solar Irradiance in South Africa using Machine Learning Models” for the Master of Science degree in Statistics at the University of Venda, hereby submitted by me, has not been submitted for any degree at this or any other university, that it is my own work in design and in execution, and that all reference material contained therein has been duly acknowledged.

Signature: Date: 12 August 2020

# Dedication

*To my mom Livhuwani, my grandmother Agnes, and my sister Adivhaho*

# Acknowledgment

I would like to thank my supervisor Dr Caston Sigauke and my co-supervisor Mr Rendani Mbuvha who guided, supported and motivated me throughout my research project. Their availability and critiques gave me courage and strength to do my research project.

I am also grateful to the University of Venda for accepting me as their student, the DST-CSIR National e-Science Postgraduate Teaching and Training Platform (NEPTTP) for the financial support. My special gratitude also goes to my colleagues and friends, namely Maduvha, Lucky, Thakhani, and Masala for the encouragement and help they offered. I am really indebted to you all.

Special thanks goes to my family, who have always been there for me, through thick and thin. It is because of their invaluable support and assistance that I always have the strength and courage to pursue my dreams. I sincerely thank the Southern African Universities Radiometric Network (SAURAN) for helping with the provision of the data used in this project.

Finally, to the Father, the Creator and Sustainer of all things, the ultimate Father of fathers, thank you for life, knowledge, and wisdom.

# Table of contents

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| ANFIS | Adaptive Neuro-Fuzzy Inference System |
| ANN | Artificial Neural Network |
| AAKR | Auto Associative Kernel Regression |
| ARIMA | Auto-Regressive Integrated Moving Average |
| CARDS | Coupled Auto-Regressive and Dynamical System |
| DBN | Deep Belief Networks |
| DCE | Decomposition-Clustering-Ensemble |
| DNN | Deep Neural Network |
| EEMD | Ensemble Empirical Mode Decomposition |
| EMD | Empirical Mode Decomposition |
| ERBF | Exponential Radial Basis Function |
| ESRAM | European Solar Radiation Atlas Model |
| FFNN | Feed Forward Neural Network |
| GCV | Generalised Cross Validation |
| GHI | Global Horizontal Irradiance |
| HGSR | Horizontal Global Solar Radiation |
| LASSO | Least Absolute Shrinkage and Selection Operator |
| LSSVR | Least Square Support Vector Regression |
| LSTM | Long Short-Term Memory |
| KNN | K-Nearest Neighbor |
| MAE | Mean Absolute Error |
| MAPE | Mean Absolute Percentage Error |
| MLPNN | Multi-Layer Perceptron Neural Network |
| NWP | Numerical Weather Prediction |
| PCR | Principal Component Regression |
| PINC | Prediction Interval with Nominal Confidence |
| PINAW | Prediction Interval Normalised Average Width |
| PIW | Prediction Interval Width |

| | |
|---|---|
| PLAQR | Partially Linear Additive Quantile Regression |
| QRA | Quantile Regression Averaging |
| RBF | Radial Basis Function |
| RBFNN | Radial Basis Function Neural Network |
| RMSE | Root Mean Squared Error |
| RNN | Recurrent Neural Network |
| SAURAN | Southern African Universities Radiometric Network |
| SVD | Singular Value Decomposition |
| SVM | Support Vector Machine |
| SVR | Support Vector Regression |
| UPR | University of Pretoria radiometric |
| $W/m^2$ | Watts Per Square Meter |

# Chapter 1

# Introduction

As the global population and levels of industrialisation continue to increase rapidly, fossil fuels which are used for electricity generation deplete rapidly as well (Zendehboudi et al., 2018). The continuous use of fossil fuels for electricity generation also continues to cause environmental problems such as global warming (Sun et al., 2018). Hence, there is a need for researchers to focus extensively on renewable energy sources such as solar, wind, waves, geothermal heat and others. Renewable energy sources are environmentally friendly, clean and inexhaustible (Mohammadi et al., 2016).

Solar energy is one of the most important forms of renewable energy that can contribute in the electrical grid to confront current environmental and energy challenges (Zhandire, 2017). However, the integration of solar energy into the electrical grid needs accurate forecasts for good and effective management of the electrical grid (Cristaldi et al., 2017). Electricity utility decision makers face a challenge of balancing demand and supply of electricity in a cost effective way which also favours future economic prosperity and environmental security. Solar irradiance forecasting is important in different applications such as agriculture, seawater desalination, medical studies and photovoltaic applications (Rezrazi et al., 2016).

## 1.1 Background

Studies on solar energy were first initiated by Liu and Jordan (1960). Their studies focused on the relationship between daily diffuse and global irradiance components on clear days on a horizontal surface, with the measurements from 98 sites in the US and Canada. Since then, researchers have been giving attention to solar irradiance with many modelling techniques being employed. Solar irradiance forecasting mainly consists of physical models and statistical data-driven models (Yang et al., 2013). Physical models are based on numerical weather predictions (NWP) for solar irradiance forecasting. According to Sun et al. (2018), solar irradiance forecasting techniques can be classified into three groups, traditional mathematical statistics, numerical weather forecasting and machine learning. Researchers have recently paid attention to machine learning techniques such as artificial neural networks (ANN) (Deb et al., 2018), and support vector machines (SVM) (Fan et al., 2018) in forecasting solar irradiance.

According to literature in solar irradiance forecasting, different forecast horizons have been explored such as long-term (a time span of three or more years), short-term (typically less than 3 months but has a time span of up-to 1 year) and medium-term (typically 3 months to 1 year but has a time span from one to three years). Feature variables used in previous studies include meteorological variables such as temperature, humidity, precipitation, cloud cover, wind speed; geographical variables such as latitude, longitude, altitude, etc; and astronomical variables such as declination, hour angle, and zenith angle as explanatory variables (Khatib and Elmenreich, 2015). This study focuses on modelling solar irradiance using long short-term memory (LSTM) networks, support vector regression (SVR), and feed forward neural networks (FFNN) models which have not been applied on South African solar irradiance data, to the best of our knowledge. The Southern African region experience sunshine all year round, therefore, South Africa's local resource is one of the highest in the world. There has been little research in South Africa on solar irradiance forecasting and this project seeks to make a contribution in this critical area.

## 1.2 Purpose of the Study

### Aim

The aim of this research project is to compare predictive performance of LSTM networks, SVR, and FFNN models on forecasting solar irradiance at the University of Pretoria radiometric (UPR) station in South Africa.

### Objectives

The main objectives of this study are to:

i develop machine learning models for solar irradiance forecasting,

ii evaluate the predictive performance of the models,

iii combine forecasts from these models, iv forecast hourly solar irradiance,

v evaluate the accuracy of the forecasts,

vi suggest areas for future work.

## 1.3 Scope of the Study

The remainder of the project is organised as follows. In Chapter 2, a literature review on forecasting solar irradiance using the proposed machine learning models and other techniques is presented. The research methodology is discussed in Chapter 3. A discussion of the empirical results is presented in Chapter 4, and the project concludes in Chapter 5.

# Chapter 2

# Literature review

## 2.1 Introduction

This chapter provides an overview of solar irradiance forecasting studies in South Africa, as well as the summaries of some studies that have used the proposed methodology to forecast solar irradiance in different regions of the world.

## 2.2 Solar Irradiance Components

Solar irradiance can be divided into three types, namely, global horizontal irradiance (GHI), direct normal irradiance, and diffuse horizontal irradiance. This is due to solar irradiance being reflected, absorbed, scattered and transmitted as it moves from the sun in the form of electromagnetic waves or sun's rays. GHI is the sum of direct normal irradiance and diffuse horizontal irradiance with the account of zenith angle of the sun. Direct normal irradiance is the irradiance which travels straight to the earth's surface directly from the sun. The irradiance which is reflected or scattered is known as diffuse irradiance. The surface of the earth receive direct solar irradiance and the diffuse or reflected solar irradiance that appears to come from various directions over the entire sky due to atmospheric scattering (Badescu, 2014).

## 2.3 An Overview of Solar Irradiance Forecasting in South Africa

Solar irradiance forecasting in South Africa using different methods and techniques is discussed in literature. Cristaldi et al. (2017) present a hybrid approach model for solar irradiance and photovoltaic (PV) power short-term forecasting. The prediction of solar irradiance is done using physical models known as clear-sky models which estimate solar irradiance in the absence of clouds. Short-term PV forecasting is then done using an auto associative kernel regression (AAKR) technique which is usually used for fault detection. An application of the proposed model is done using a PV plant located in South Africa. The empirical results from this study show that the developed model produces accurate solar irradiance forecasts.

Zhandire (2017) proposes a solar resource index that is based on a site-specific sky conditions due to stochastic movement and evolution of clouds. Solar resourse index is proposed to differentiate persistent clear-sky conditions from persistent overcast-sky conditions. This study used data from eight different weather stations in South Africa. After applying the K-means clustering technique five classes are identified for the solar irradiance resource. The solar resource classification index is superior to probability of persistence index and fractal dimension index.

Adeala et al. (2015) present a study on prediction of global solar irradiance using multiple linear regression with an inclusion of weather parameters in addition to the traditional extraterrestrial irradiance and sunshine hours. The study is done in all nine provinces in South Africa. The study shows that an inclusion of weather parameters improves the accuracy and performance of solar irradiance models for some locations.

A more recent study is that of Mpfumali et al. (2019) who uses partially linear additive quantile regression (PLAQR) model in predicting day ahead GHI. The study uses data from the Tellerie radiometric station in the Northern Cape province. Two PLAQR models are developed, one without interaction and the other with interactions obtained using the least absolute shrinkage and selection operator (Lasso). This study also combines forecasts from individual models using the convex combination method and quantile regression averaging (QRA). The results of the study show that QRA forecast combination model is the best forecasting model compared with other individual models in the study.

## 2.4 Machine Learning Techniques

Artificial neural networks (ANNs) have been applied extensively in forecasting solar irradiance. ANNs based models already applied are feed forward ANN, cascade-forward back propagation ANN, generalised regression ANN, neurofuzzy ANN, and optimized ANN-genetic algorithm. ANNs have the ability to map relationships between input and output variables provided there is historical data to learn from (Haykin et al., 2009).

Paoli et al. (2010) present an integrated model to forecast daily solar irradiance time series using neural networks (ANNs). Multi-layer perceptron neural network (MLPNN) is used in the study for daily solar irradiance forecasting after using the seasonal index adjustment method to adjust original solar irradiance sequence. The outcome of the study shows that the mean absolute percentage error (MAPE) of the MLPNN model is lower than that of autoregressive integrated moving average (ARIMA), the Bayesian, Markov chain model and K-nearest neighbor (KNN) technique. Benmouiza and Cheknane (2013) use K-means clustering to classify input data into regions of the reconstructed phase-space which has similar characteristics, then model different groups using a non-linear autoregressive neural network, and then forecast the solar irradiance on test data by the corresponding model. The results of the study show that the clustering of the input space is an important task to interpret the behaviour of the series.

Gensler et al. (2016) study different ANN and deep neural network (DNN) architectures in the field of solar power forecast. Different models such as physical forecasting model, MLPNN, LSTM, deep belief networks (DBN), and Auto-Encoders are analyzed and compared. The results of the study show that Deep Learning algorithms have superior solar power forecasting performance compared to ANN and physical models on data from 21 solar power plants in Germany.

Shamshirband et al. (2016) present a hybrid support vector machine firefly optimization algorithm (SVM-FFA) model to estimate monthly mean horizontal global solar radiation (HGSR) at Bandar in Iran. Using the approach and long-term measured HGSR, three models are calibrated by considering different sets of meteorological parameters measured for Bandar Abbass situated in Iran. It is found that the model utilizing the combination of relative sunshine duration, difference between maximum and minimum temperatures, relative humidity, water vapour pressure, average temperature, and extraterrestrial solar irradiance shows superior performance compared to all the other approaches used. The survey results reveal that the developed SVM-FFA approach is capable of providing favourable predictions with significantly higher precision than other examined techniques.

Sun et al. (2018) propose a decomposition-clustering-ensemble (DCE) learning approach for solar irradiance forecasting. The application results show that the EEMD-LSSVR-K (K-means clustering)-LSSVR learning technique can significantly improve the predicting performance and is superior to some popular forecasting methods in terms of forecasting accuracy and robustness analysis. Sreekumar et al. (2016) presents a SVR model with fixed parameter optimization and two hyper parameter optimized SVR models, support vector regression with optimized hyper parameters using Genetic Algorithm (SVRGA) as well as Particle Swarm Optimization (SVRPSO) for short-term solar irradiance forecasting. Their study uses solar irradiance data from Chicago, USA. The results show that SVRPO outperform SVR model with fixed parameter optimization and SVRGA.

## 2.5 Non-Machine Learning Techniques

A number of alternative methods have been used to forecast solar irradiance besides machine learning techniques. Huang et al. (2013) propose a coupled autoregressive and dynamical system (CARDS) to forecast the solar irradiance. The proposed method increases the 1-hour ahead global solar irradiance forecast accuracy by 30% than general neural network or random models. Akarslan and Hocaoglu (2016) first use the multidimensional linear predictive filtering model to forecast hourly solar irradiance in different regions of Turkey, which is superior to the two-dimensional linear predictive filtering model and the traditional statistical forecasting method by means of empirical analysis.

Akarslan et al. (2018) propose five semi-empirical models to forecast hourly solar irradiance which use historical data of solar irradiance, extraterrestrial irradiance, and clearness index. These models are applied on data from three regions in Turkey. The results show that the proposed models outperform Angstrom-Prescott models based on the forecast accuracy.

Novel hybrid adaptive neuro-fuzzy inference system (ANFIS) models with particle swarm optimization, genetic algorithm and differential evolution are proposed by Halabi et al. (2018). The proposed ANFIS model with particle swarm optimization outperformed other hybrid ANFIS models on a Malaysian dataset. Li et al. (2018) proposed a rolling prediction model combining empirical mode decomposition (EMD) and ANN techniques which uses historical solar irradiance data only. The proposed methodology is applied to a Chinese dataset. The results reveal that the proposed method is comparable to previous research studies.

## 2.6 Conclusions from Literature

The use of ANNs and SVR on solar irradiance forecasting has been discussed in detail under literature. This research project seeks to investigate the application of ANNs and SVR in forecasting solar irradiance in South Africa.

# Chapter 3

# Methodology

## 3.1 Introduction

This chapter discusses the methods and techniques that will be used in this study to forecast hourly global solar irradiance. The feed forward neural networks (FFNN), long short-term memory (LSTM) network, and support vector regression (SVR) are discussed. This chapter will also discuss techniques that are going to be applied in model selection and in assessing forecast accuracy.

## 3.2 Feed Forward Neural Networks

A feed forward neural network (FFNN) is a type of artificial neural network (ANN) in which connections between the nodes do not form a cycle or a loop. The study of this technique was first initiated by McCulloch and Pitts (1943) who created a computational model of the concept. Different researchers have further expanded the concept of ANN to cover many features (Rosenblatt, 1958; Minsky and Papert, 1990). FFNN works by feeding input data in one end which is then processed by the network and comes out as output in the other end. Information flows in the forward direction only.

### 3.2.1 Single-Layer Perceptron

Single-layer perceptron is the simplest type of a neural network which has a single layer of output nodes (McCulloch and Pitts, 1943). A single-layer neural network can be described mathematically as follows:

$$y_k = \left( \sum_{i=0}^{D} \omega_i x_i \right) \tag{3.1}$$

where $y_k$ is the output, $g(\cdot)$ is an activation function, $x_i$ is input and $\omega_i$ represents the corresponding weight for $x_i$. Single-layer neural networks are not usually used in practice, but help in understanding the basic concept of neural networks.

### 3.2.2 Multi-Layer Perceptron

Multi-layer perceptron consists of multiple layers of computational units, usually interconnected in a feed-forward way (McCulloch and Pitts, 1943). A multi-layer neural network with one hidden layer can be written as

$$y_k = h\left( \sum_{j=0}^{M} \omega_{kj}^{(2)} g(a_j) \right) \tag{3.2}$$

where,

$$a_j = g\Big( \sum_{i=0}^{D} \omega_{ij}^{(1)} x_i \Big). \tag{3.3}$$

The multi-layer neural network is very similar to single-layer neural network except that multi-layer neural network's output of the inner layer is again multiplied by a new weight vector and wrapped in an activation function.

## 3.3 Long Short-Term Memory networks

The LSTM (long short-term memory) network is a type of recurrent neural network (RNN). RNN is used when dealing with sequential data, like time series data. The LSTM network was originally introduced by Hochreiter and Schmidhuber (1997) in a paper titled *Long Short-Term Memory.* The LSTM network is different from regular RNN due to the reason that it consists of LSTM blocks instead of nodes. One LSTM block can be represented by the following system of equations,

$$\begin{aligned} f_t &= g(W_t.[x_t, h_{t-1} + b_f]) \\ i_t &= g(W_i.[x_t, h_{t-1} + b_i]) \\ o_t &= g(W_o.[x_t, h_{t-1} + b_o]) \\ c_t &= f_t c_{t-1} + i_t \tanh(g(W_c.[x_t, h_{t-1} + b_c])) \\ h_t &= o_t \odot \tanh(c_t), \end{aligned} \tag{3.4}$$

where $g(\cdot)$ is the sigmoid function, $\tanh(\cdot)$ is the hyperbolic tangent, $x_t$ is the input vector, $h_t$ is the output vector, $c_t$ is a cell state vector, $W$ are weights and $b$ are biases, $f_t$, $i_t$, and $o_t$ are gates of the block.

## 3.4 Support Vector Regression

Support vector regression (SVR) is based on support vector machine (SVM) which is a supervised machine learning technique which involves statistical learning theory and the principle of structural risk minimization. SVR was introduced by Drucker et al. (1997) and was extended from SVM model. There are different basic kernel functions that are used in SVM models, which can be classified as polynomial (Poly), Gaussian kernel, exponential radial basis function (ERBF), radial basis function (RBF), sigmoid and linear (Zendehboudi et al., 2018). The SVR works by mapping the input space into a high-dimensional feature space and constructs the linear regression in it which can be expressed as

$$f(x) = \omega\phi(x) + b, \tag{3.5}$$

where $\omega$ is the weight vector, $\phi(x)$ maps inputs $x$ into a high dimensional feature space that is nonlinearly mapped from the input space $x$, and $b$ is the bias term.

To calculate the coefficients $\omega$ and $b$ it is required to reduce the regularised risk function which can be expressed as

$$\frac{1}{2}\left\|\omega\right\|^2 + C\frac{1}{l}\sum_{i=1}^{l} L_\epsilon(y_i, f(x_i)), \tag{3.6}$$

where $\left\|\omega\right\|^2$ is a regularised term which maintains the function capacity. $C$ is a cost error. The empirical term from the second term in equation 3.6 can be defined as

$$L_{\epsilon}(y_i, f(x_i)) = \{|y_i = f(x_i)| = \epsilon, |y_i = f(x_i)| \geq \epsilon\}. \tag{3.7}$$

Equation 3.6 expressed the transformation of the primal objective function in order to get the values of $\omega$ and $b$ by introducing the positive slack variables $\xi_i(^*)$.

$$\text{minimise} \quad \frac{1}{2}\left\|\omega\right\|^2 + C\frac{1}{l}\sum_{i=1}^{l}(\epsilon_i + \xi_i^*) \tag{3.8}$$

subject to

$$\alpha(x) = \begin{cases} y_i - \langle \omega, x_i \rangle - b \leq \epsilon + \xi_i \\ \langle \omega, x_i \rangle + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases}$$

The optimization problem expressed in equation 3.8 has to be transformed into its dual formulation by using the Lagrange multipliers to solve it in a more efficient way as follows:

$$\begin{aligned} L = \frac{1}{2}\left\|\omega\right\|^2 + C\frac{1}{l}\sum_{i=1}^{l}(\epsilon_i + \xi_i^*) - \sum_{i=1}^{l}(\eta_i\epsilon_i + \eta_i\xi_i^*) - \sum_{i=1}^{l} a_i(\epsilon + \xi_i + y_i + \omega.\phi(x_i) + b) \\ - \sum_{i=1}^{l} a_i^*(\epsilon + \xi_i^* - y_i - \omega.\phi(x_i) + b), \end{aligned} \tag{3.9}$$

where $L$ is the Lagrange and $\eta_i, \eta_i^*$ are the Lagrange multipliers. Hence the dual variables in equation 3.6 have to satisfy positive constraints, $\eta_i^*, a_i^* \geq 0$. The resulting SVR model can be expressed as follows:

$$f(x) = \sum_{i=1}^{l}(a - a_i^*)\phi(x_i)\phi(x) + b \tag{3.10}$$

where $a_i$ and $a_i^*$ are Lagrange multipliers. SVR is known for being effective in high-dimensional space and effective in memory space since it relies on using support vectors.

## 3.5 Principal Component Regression

Principal Component Regression (PCR) is a regression analysis that is used to deal with multiple regression data that has multicollinearity (Liu et al., 2003). PCR model will be used in this study as a benchmark model. Multicollinearity occurs when one predictor variable in a multiple regression model can predicted with other variables linearly. It leads to least squares estimates which have large variances, which means they maybe distant from their true values. A multiple regression model is given by

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \tag{3.11}$$

where $\mathbf{Y}$ is a vector of observed values, $\mathbf{X}$ is a matrix of explanatory variables, $\beta$ is a parameter vector, and $\varepsilon$ is vector of error terms. The least squares solution of equation 3.11 is given by

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}. \tag{3.12}$$

To get the first $b$ Principal Components (PCs), we approximate the matrix $\mathbf{X}$ using singular value decomposition (SVD):

$$\mathbf{X} = \tilde{\mathbf{X}}_{(b)} + \varepsilon_{\mathbf{X}} = (U_{(b)}D_{(b)})V_b^T + \varepsilon_{\mathbf{X}} = \mathbf{T}_{(b)}\mathbf{P}_{(b)}^T + \varepsilon_{\mathbf{X}}, \tag{3.13}$$

where $\mathbf{T}$ represents orthogonal scores and $\mathbf{P}$ loadings. Now regressing $Y$ on the scores leads to

$$\hat{\beta} = \mathbf{P}(\mathbf{T}^T\mathbf{T})^{-1}\mathbf{T}^T\mathbf{Y}. \tag{3.14}$$

## 3.6 Variable Selection

Variable Selection involves selection of feature variables that explains a target variable thereby reducing number of feature variables. This process is beneficial in terms of avoiding overfitting, making a model easier to interpret, and reduces in computational time. There are many variable selection methods, however, in this study we use Lasso (least absolute shrinkage and selection operator) via hierarchical interactions (Bien et al., 2013). Lasso via hierarchical interactions considers pairwise hierarchical interactions only, however, it can be extended to higher order interactions. We assume a regression model with response variable Y and predictors X$_1$,...,X$_p$ with pairwise interactions between these predictors. Lasso hierarchical model is given as

$$\mathrm{Y} = \beta_0 + \sum_j \beta_j X_j + \frac{1}{2}\sum_{j\neq k} \Theta_{jk} X_j X_k + \varepsilon, \tag{3.15}$$

where $\varepsilon \sim N(0, \sigma^2)$, $\beta \in \mathbb{R}^p$, and $\Theta \in \mathbb{R}^{p\times p}$. There are two categories of hierarchical restrictions, which are strong and weak hierarchy.

$$\begin{aligned} Strong\ hierarchy : \hat{\Theta}_{jk} &\neq 0 \Rightarrow \hat{\beta}_j \neq 0\ and\ \hat{\beta}_k \neq 0 \\ Weak\ hierarchy : \hat{\Theta}_{jk} &\neq 0 \Rightarrow \hat{\beta}_j \neq 0\ or\ \hat{\beta}_k \neq 0 \end{aligned}$$

The major advantage of Lasso via hierarchical interactions is that it leads to simpler and more interpretable models that involve only a subset of the predictors.

## 3.7 Forecast Combination

Forecast combination is a method used to combine forecasts from different fitted models with a purpose of improving forecast accuracy (Gaba et al., 2017). There are many forecast combination methods, but this research will only focus on quantile regression averaging (QRA) and convex combination method.

### 3.7.1 Quantile Regression Averaging

Quantile regression averaging (QRA) was first initiated by Maciejowska et al. (2016). QRA treats forecasts from different models as independent variables and actual observations as a dependent variable (global horizontal irradiance). Let $\hat{y}_{t,\tau}$ be hourly global horizontal irradiance, $K$ be methods used to forecast the next $m$ observations, i.e. $m$ is the total number of forecasts. The forest combination, $\hat{y}^{QRA}_{t,\tau}$, is given by

$$\hat{y}^{QRA}_{t,\tau} = \beta_0 + \sum_{k=1}^{K} \beta_{t,k}\hat{y}_{t,k} + \varepsilon_{t,\tau}, \quad \tau \in (0,1), \quad t = 1, ..., m, \tag{3.16}$$

where $\hat{y}_{t,k}$ represents predictions from method $k$, $\hat{y}^{QRA}_{t,\tau}$ is the combined forecasts, and $\varepsilon_{t,\tau}$ is the error term. QRA aims to minimise

$$\underset{\beta}{\operatorname{argmin}} \sum_{t=1}^{n} \rho_\tau(\hat{y}^{QRA}_{t,\tau} - \beta_0 - \sum_{k=1}^{K} \beta_{t,k}\hat{y}_{t,k}). \tag{3.17}$$

In matrix form, we have

$$\underset{\beta \in IR}{\operatorname{argmin}} \sum_{t=1}^{n} \rho_\tau(\hat{y}^{QRA}_{t} - x_t^T\beta). \tag{3.18}$$

### 3.7.2 Convex Combination

Convex combination method computes the sequence of instantaneous losses suffered by the predictions from the experts (models) using a loss function

(Gaillard and Goude, 2015). The loss function can be based on square, absolute, percentage, or pinball loss. The combined forecasts will be compared with forecasts from each model using the equation given as

$$\hat{y}^{c}_{t,\tau} = \sum_{m=1}^{M} \omega_{mt} \hat{y}_{mt,\tau}, \tag{3.19}$$

where $\omega_{mt}$ is the weight given to forecast $m$.

## 3.8 Prediction Intervals

The prediction interval widths (PIWs) for every model, $M_j, j = 1, ..., k$, are denoted as $\text{PIW}_{ij}, i = 1, ..., m, j = 1, ..., k$. $\text{PIW}_{ij}$ is calculated as

$$\text{PIW}_{ij} = \text{UL}_{ij} - \text{LL}_{ij}, \tag{3.20}$$

where $\text{UL}_{ij}$ and $\text{LL}_{ij}$ are the upper and lower limits of the prediction interval, respectively. Probability density plots and box and whisker plots will be used in this study to find the model which yields narrower PIWs.

## Evaluation of Prediction Intervals

A prediction interval with nominal confidence (PINC) of $100(1 - \alpha)\%$ is defined as the probability that the forecast $\hat{y}_{t,\tau}$ lies in the prediction interval $(\text{LL}_{ij}, \text{UL}_{ij})$ (Sun et al., 2017). PINC is given by

$$\text{PINC} = P(\hat{y}_{t,\tau} \in (LL_{ij}, UL_{ij}) = 100(1 - \alpha)\%. \tag{3.21}$$

There are many indices used to evaluate the reliability of prediction intervals, however, in this study we use the prediction interval coverage probability

(PICP), the prediction interval normalised average width (PINAW) (Sun et al., 2017). PICP is given by

$$\text{PICP} = \frac{1}{m}\sum_{i=1}^{m} I_{ij}, \tag{3.22}$$

where $m$ is the number of forecasts and $I$ is a binary variable given by,

$$I_{ij} = \begin{cases} 1, & \text{if} \quad y_i \in (LL_{ij}, UL_{ij}) \\ 0, & \text{otherwise.} \end{cases} \tag{3.23}$$

PINAW is another index used to evaluate the reliability of prediction intervals and is given by

$$\text{PINAW} = \frac{1}{m(\max(y_{ij}) - \min(y_{ij}))}\sum_{i=1}^{m}(UL_{ij} - LL_{ij}), j = 1, ..., k. \tag{3.24}$$

## 3.9 Evaluation of Forecasts

The mean absolute error (MAE) and root mean squared error (RMSE) are going to be used to evaluate the accuracy of our forecasts. The equations of the above measures are respectively given by

$$\text{MAE} = \frac{1}{m}\sum_{t=1}^{m} \left| y_t - \hat{y}_t \right|, \tag{3.25}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{m}(y_t - \hat{y}_t)^2}{m}}, \tag{3.26}$$

where $\hat{y}_t$ are predicted values by the model, $y_t$ are the values actually observed, and $m$ is the number of predictions.

## 3.10 Data and Features

This study uses data obtained from the Southern African Universities Radiometric Network (SAURAN) database, accessible at http://sauran.net. We focus our work on only one radiometric station found at an inland region in Pretoria, South Africa. The station of focus is the University of Pretoria radiometric (UPR) station. The exact location of the station is at 28°13′42.924″E and 25°45′11.088″S as shown in Figure 3.1. This research project seeks to model hourly solar irradiance (global horizontal irradiance (GHI)) using independent variables such as temperature, wind speed, relative humidity, barometric pressure, wind direction standard deviation, rainfall, hour, month, a non-linear trend, and a lagged hourly solar irradiance at lags 1 and 2.

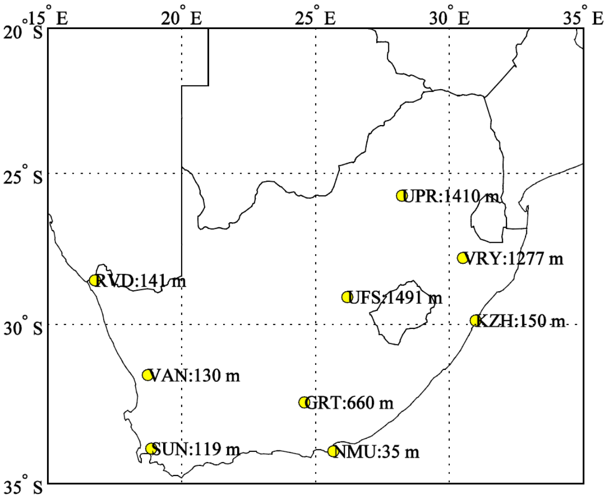


Figure 3.1: Map showing the location and altitude of the University of Pretoria radiometric (UPR) station with some selected stations in South Africa (Zhandire, 2017).

The data used is hourly solar irradiance from 1 January 2014 to 31 December 2018 with 20056 observations with hours from 7:00 AM to 5:00 PM considered to be sunshine hours. The UPR data set is split into training data, 1 January 2014 to 31 December 2017, i.e., $n_1 = 16044$ and testing data, from 1 January 2017 to 31 December 2018, i.e., $n_2 = 4012$, which is 20% of the total number of observations.

## 3.11 Computational Tools

The software packages that are used for data analysis in this study are R and Python. The FFNN and LSTM networks algorithm are implemented in this study using the Keras deep learning package (https://keras.io/). SVR is implemented using sklearn python package.

# Chapter 4

# Data Analysis

## 4.1 Introduction

This chapter presents data analysis and discussion using the methodology discussed in Chapter 3. The predictive performance of long short-term memory (LSTM) networks, support vector regression (SVR), and feed forward neural networks (FFNN) are compared in this chapter for forecasting hourly solar irradiance.

## 4.2 Exploratory Data Analysis

The summary statistics of hourly solar irradiance for the sampling period January 2014 to December 2018 is given in Table 4.1 for the UPR radiometric station. The distribution of hourly solar irradiance is not normally distributed since it is skewed to the right and platykurtic as shown by the skewness value of 0.089 and a kurtosis value of -0.997 given in Table 4.1. Figure 4.1 shows the time series plot of hourly solar irradiance together with density, normal quantile to quantile (QQ) and box plots which all show that

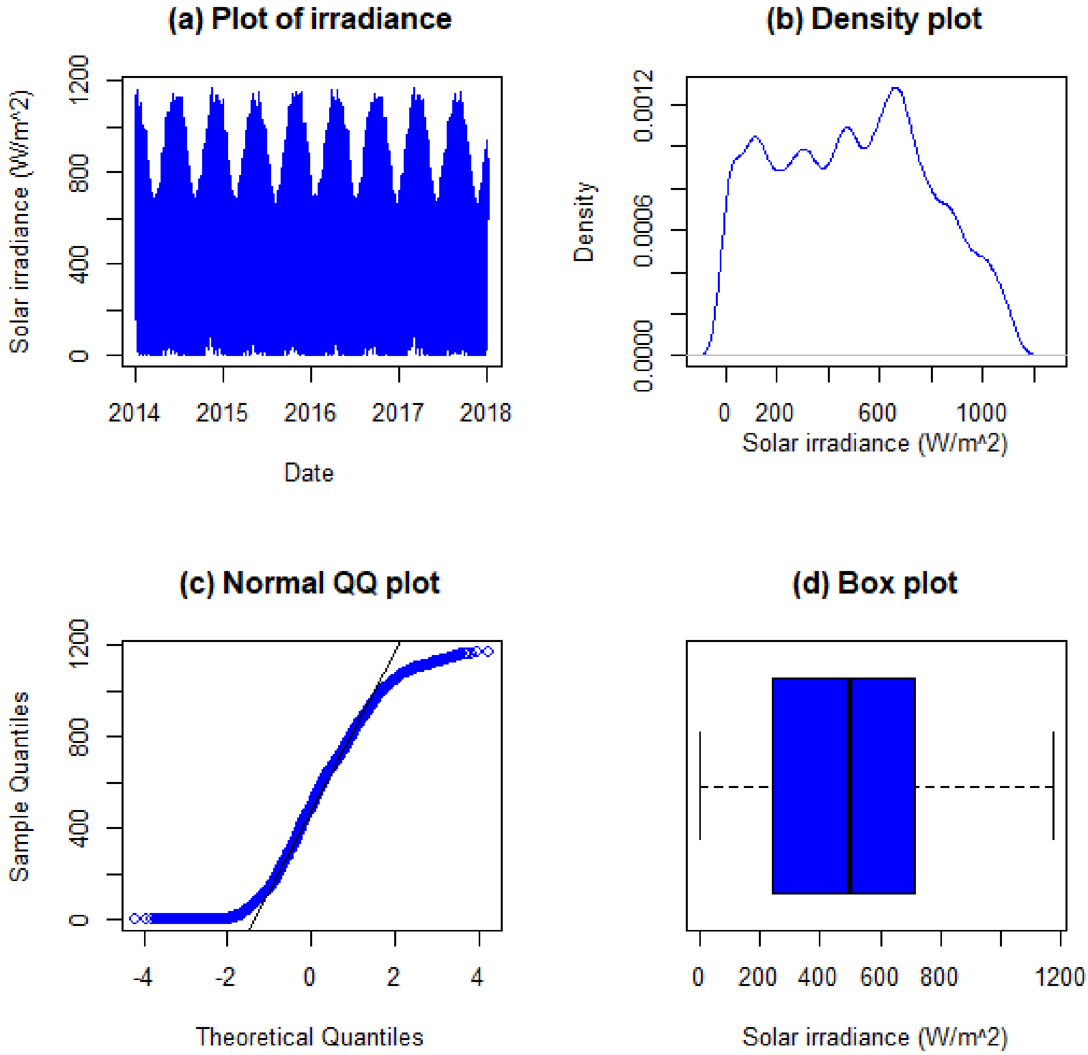


Figure 4.1: Diagnostic plots for hourly solar irradiance $(W/m^2)$.

the data is not normally distributed.

Table 4.2 summarises the weather variables considered for this study and their ranges.

Table 4.1: Descriptive statistics for hourly solar irradiance ($W/m^2$).

| Min | Median | Mean | Max | St.Dev. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| 0.0 | 496.7 | 496.6 | 1173.9 | 298.760 | 0.089 | -0.997 |

Table 4.2: Weather variables and their value ranges.

| Weather variable | Value range |
|---|---|
| Temperature ($°C$) | 1.95 - 37.28 |
| Barametric pressure(mbar) | 814.00 - 878.00 |
| Relative humidity (%) | 5.09 - 97.90 |
| Rainfall (mm) | 0.00 - 21.84 |
| Wind speed (m/s) | 0.01 - 9.93 |
| Wind direction (°) | 0.05 - 360.00 |
| WD standard deviation (°) | 6.72 - 76.51 |

The non-linear trend values used in this project are obtained or extracted by fitting a cubic smoothing spline function which is given by:

$$\pi(t) = \sum_{t=0}^{n} (y_t - f(t))^2 + \lambda \int \{f''(t)\}^2 dt \qquad (4.1)$$

where $\lambda$ is a smoothing parameter which is estimated using the generalised cross validation (GCV) criterion. Figure 4.2 shows the non-linear trend, cubic smoothing spline fitted with an estimated $\lambda$ value. The fitted non-linear trend values are extracted and used to model solar irradiance.

Figure 4.3 shows box plots of the distribution of hourly irradiance data from 2014 to 2018 for each month and hour of the day, respectively. It can be clearly seen that solar irradiance shows seasonality, since its values are low during the winter season and are high during the summer season. The solar irradiance series is also highly correlated with hour of the day since it tends

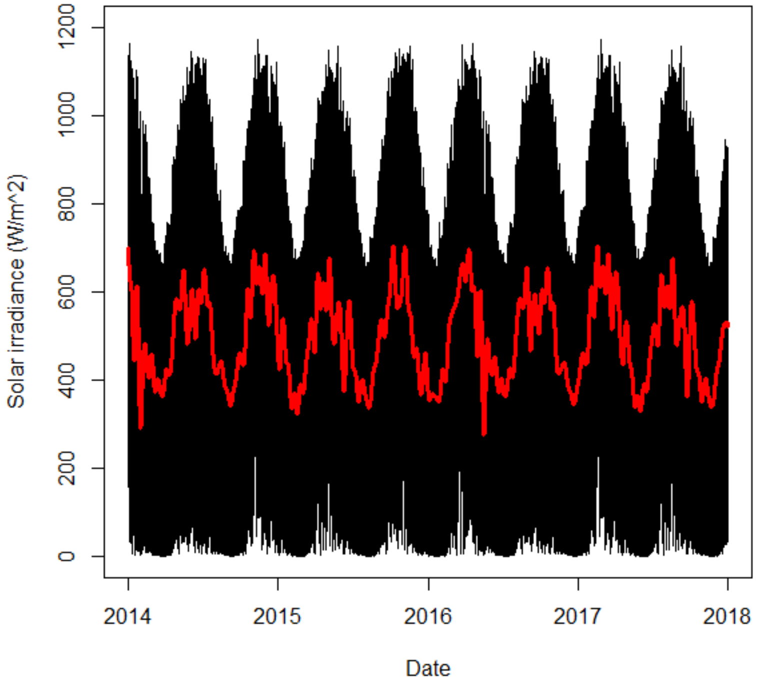


Figure 4.2: Plot of hourly solar irradiance from 1 January 2014 to 31 December 2018 superimposed with a fitted cubic smoothing spline trend (non-linear trend).

to be low in the morning and increases towards the peak in the afternoon and decreases towards the evening.

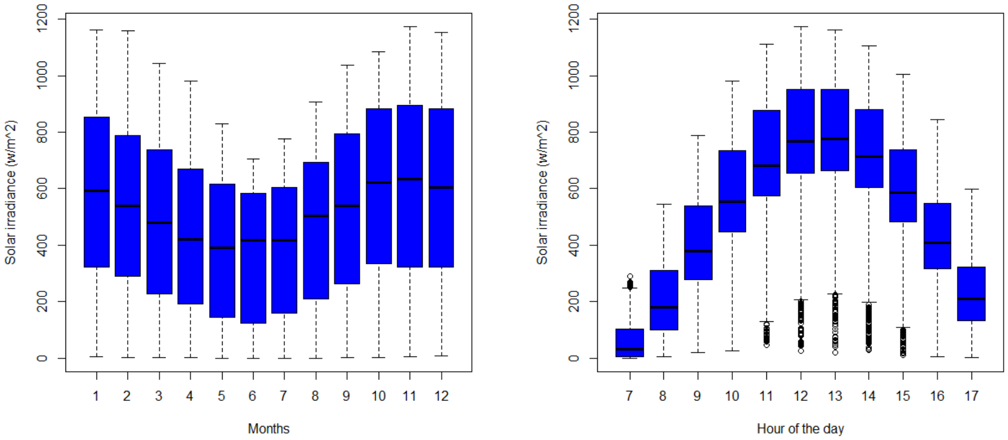


Figure 4.3: Distribution of monthly and hourly solar irradiance data.

## Variable selection using Lasso

Different weather variables are recorded by SAURAN at different radiometric stations. Variable selection is done in this study using the Least Absolute Shrinkage and Selection Operator (Lasso) in order to remain with significant variables in predicting global solar irradiance. Figure 4.4 shows the importance of 13 weather variables in predicting global solar irradiance as assessed by Lasso. Wind direction (Wd) and day of the month have the least significance in forecasting solar irradiance and are therefore not considered in the modelling stage.

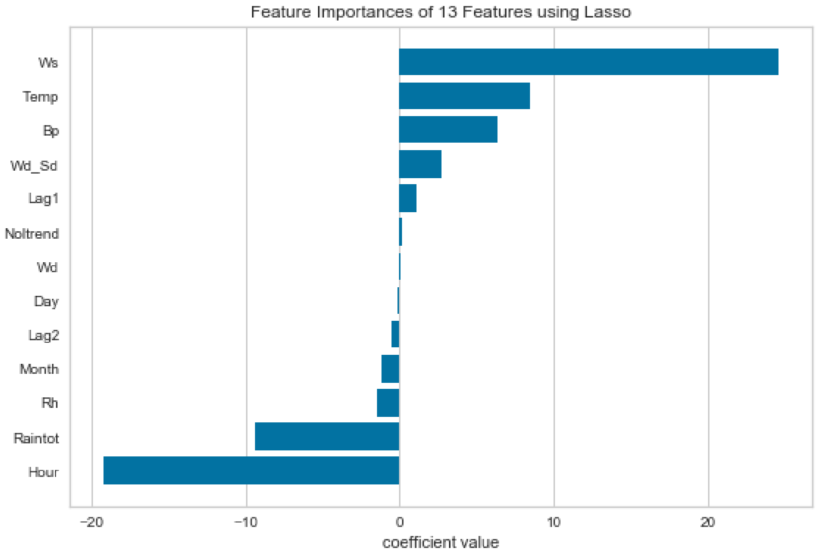


Figure 4.4: Variable selection using Lasso.

## 4.3 Machine Learning Models

The models considered are M2 (LSTM), M3 (SVR), and M4 (FFNN) which are all machine learning models and M1 (PCR) which is a benchmark model in this study. RMSE and MAE are normally used for forecast evaluation and are used in this study.

Table 4.3 shows a comparative analysis of the fitted machine learning models, together with the benchmark model. All machine learning models outperform the benchmark model on both training set and testing set. The M4 (FFNN) model has the least RMSE (83.087) and MAE (51.237), showing that it is the best fitting model according to the summary of the error measures for evaluation of the models on the training set.

Table 4.3 also summarises the error measures for evaluation of the models on the testing set. From Table 4.3 model M4 (FFNN) has the least RMSE (88.326) and MAE (53.719) among the three machine learning models. This means that model M4 is the best forecasting model.

Table 4.3: Comparative analysis of the fitted machine learning models.

| **Evaluation of the models on the training set** | | | | |
|---|---|---|---|---|
| | M1 (PCR) | M2 (LSTM) | M3 (SVR) | M4 (FFNN) |
| RMSE | 130.241 | 91.884 | 95.251 | 83.087 |
| MAE | 94.346 | 55.878 | 68.705 | 51.237 |
| **Evaluation of the models on the testing set** | | | | |
| | M1 (PCR) | M2 (LSTM) | M3 (SVR) | M4 (FFNN) |
| RMSE | 129.449 | 94.131 | 99.302 | 88.326 |
| MAE | 94.554 | 56.291 | 72.606 | 53.719 |

The graphical plot of the out of sample forecasts for the models M1, M2, M3 and M4 for the period January 2018 to December 2018 are given in Figure 4.5. The test set has 4012 observations. M2 and M4 forecasts appear to be the same. M1 and M3 forecasts differ with other forecast models, especially at the lower tails.

The probability densities of hourly solar irradiance and the forecasted values for the period January 2018 to December 2018 for machine learning models M2, M3, M4, are shown in Figure 4.6, including that of M1. The forecasts from the three models appear to have some slight difference with the actual observations.

Figures 4.7, 4.8, and 4.9 show the true solar irradiance values and their forecast results for M2 (LSTM), M3 (SVR), and M4 (FFNN), respectively, for selected days. Each of the figures show forecasts for the first day of four seasons in South Africa which are: autumn (March-May), winter (June-August), spring (September-November), and summer (December-February). The selected days are March 1 (autumn), June 1 (winter), September 1 (spring), and December 1 (summer), all in 2018. The forecasts from the three models appear to be fairly close with the actual observations, especially on the chosen spring and summer days.

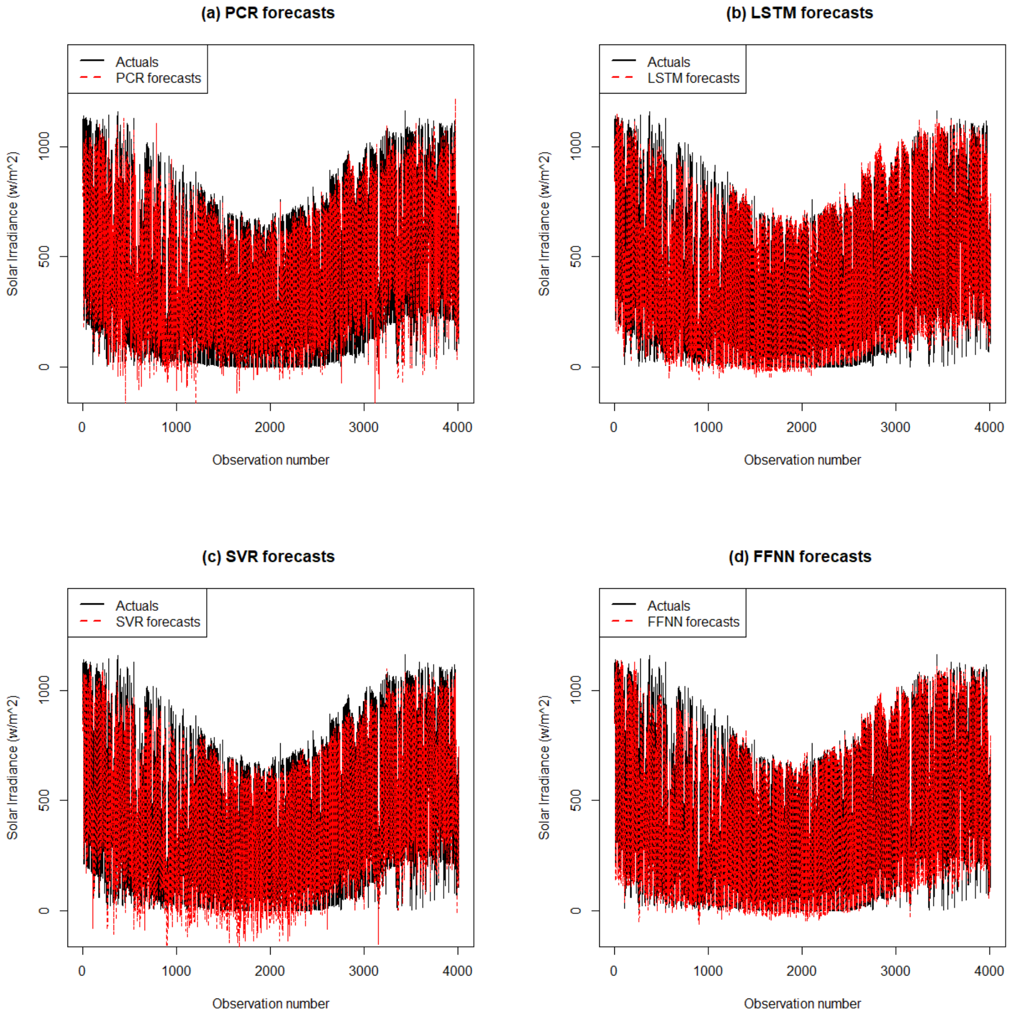
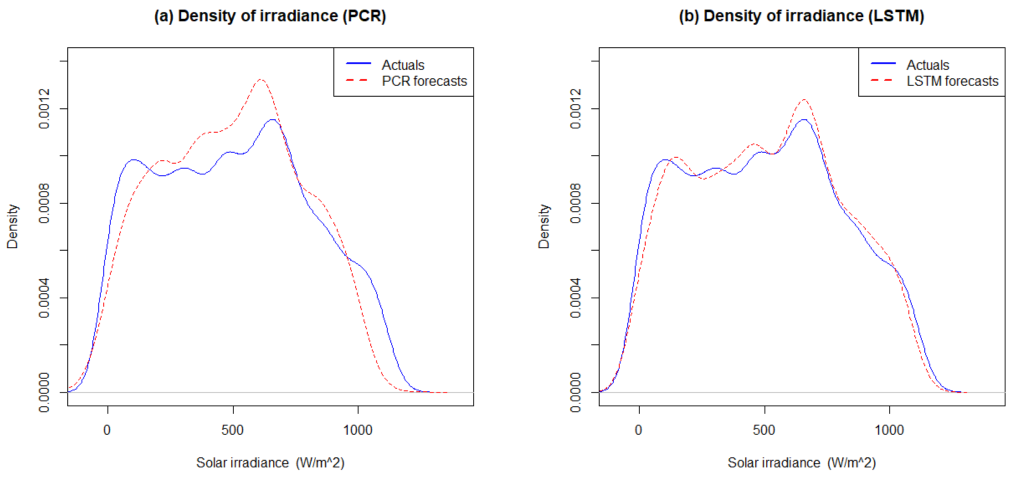


Figure 4.5: Graphical plot of (a) **Top left panel**: Forecasts (dashed line) using model M1 (PCR) and actual hourly irradiance (solid line) (b) **Top right panel**: Forecasts (dashed line) using model M2 (LSTM) and actual hourly irradiance (solid line) and (c) **Bottom left panel**: Forecasts (dashed line) using model M3 (SVR) and actual hourly irradiance (solid line) (d) **Bottom right panel**: Forecasts (dashed line) using model M4 (FFNN) and actual hourly irradiance (solid line).



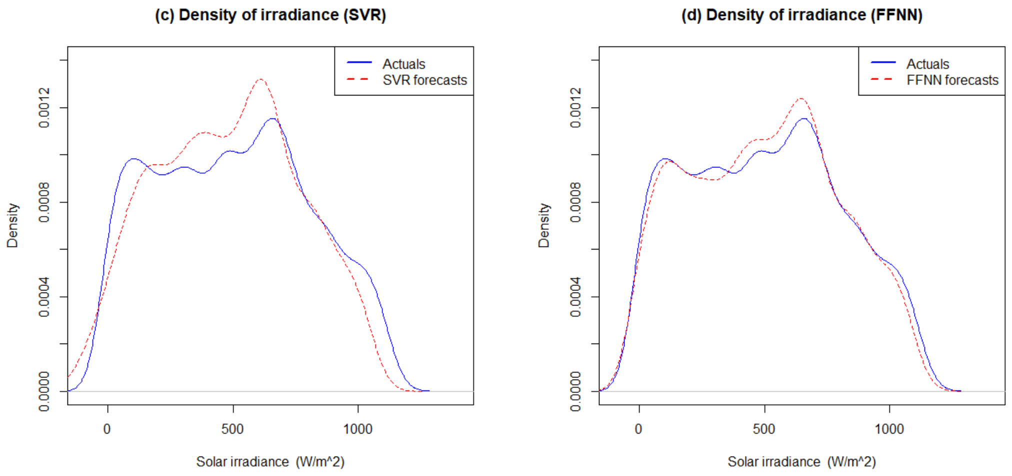


Figure 4.6: Probability densities of (a) **Top left panel**: Forecasts (dashed line) using model M1 (PCR) and actual hourly irradiance (solid line) (b) **Top right panel**: Forecasts (dashed line) using model M2 (LSTM) and actual hourly irradiance (solid line) and (c) **Bottom left panel**: Forecasts (dashed line) using model M3 (SVR) and actual hourly irradiance (solid line) (d) **Bottom right panel**: Forecasts (dashed line) using model M4 (FFNN) and actual hourly irradiance (solid line).

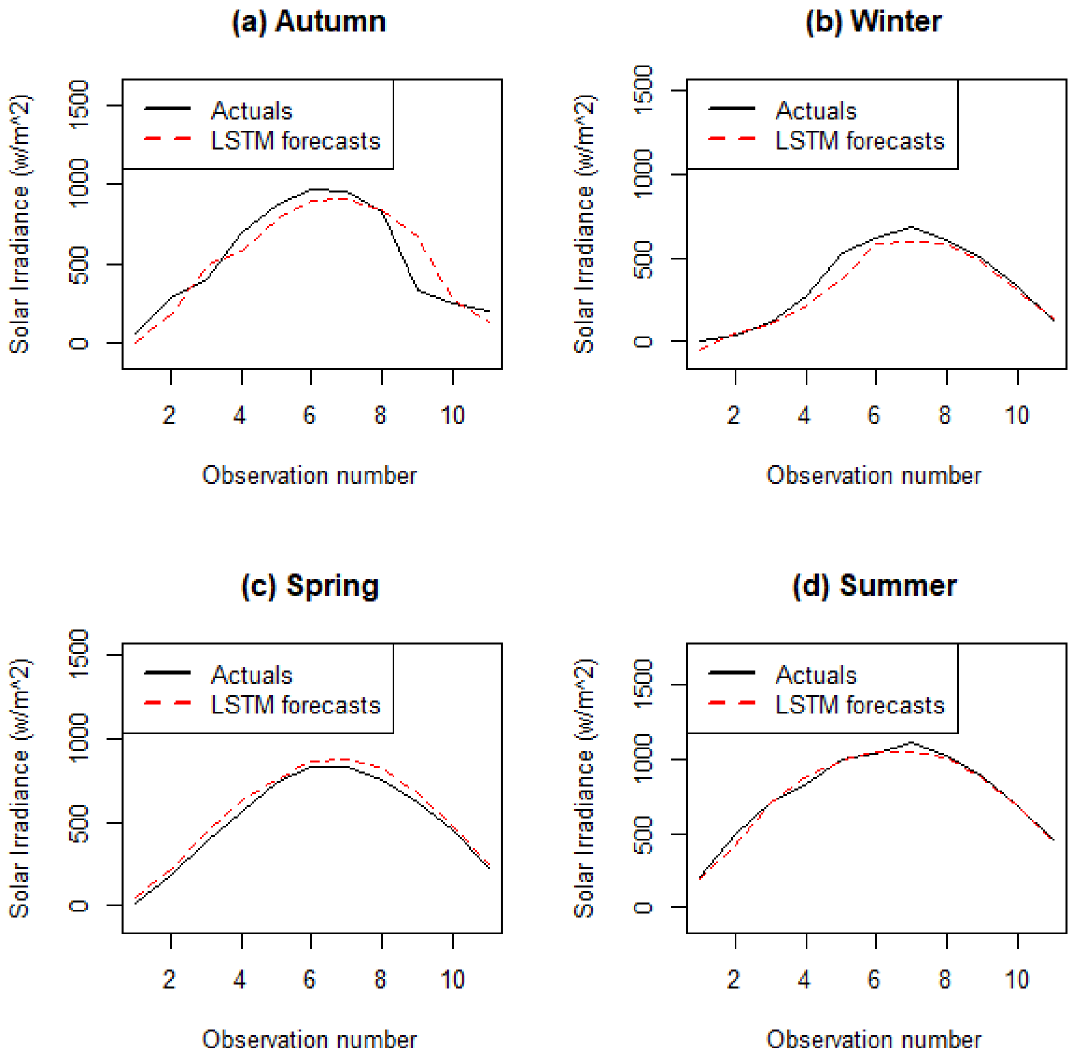


Figure 4.7: Graphical plot of forecasts (dashed line) using model M2 (LSTM) and actual hourly irradiance (solid line) for the first day of season in (a) **Top left panel**: Autumn (1 March 2018)and (b) **Top right panel**: Winter (1 June 2018) (c) **Bottom left panel**: Spring (1 September 2019) and (d) **Bottom right panel**: Summer (1 December 2019).

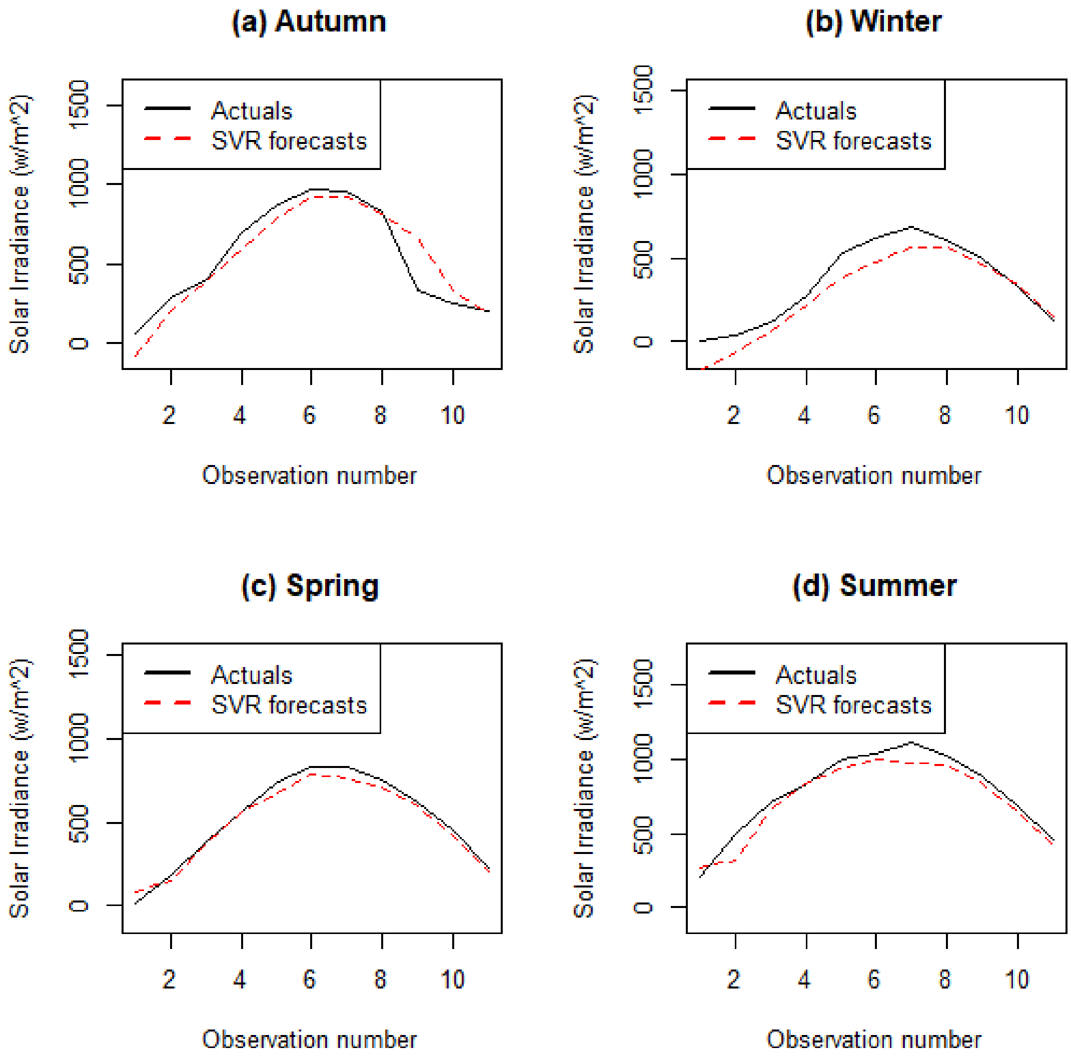


Figure 4.8: Graphical plot of forecasts (dashed line) using model M3 (SVR) and actual hourly irradiance (solid line) for the first day of season in (a) **Top left panel**: Autumn (1 March 2018) and (b) **Top right panel**: Winter (1 June 2018) (c) **Bottom left panel**: Spring (1 September 2019) and (d) **Bottom right panel**: Summer (1 December 2019).

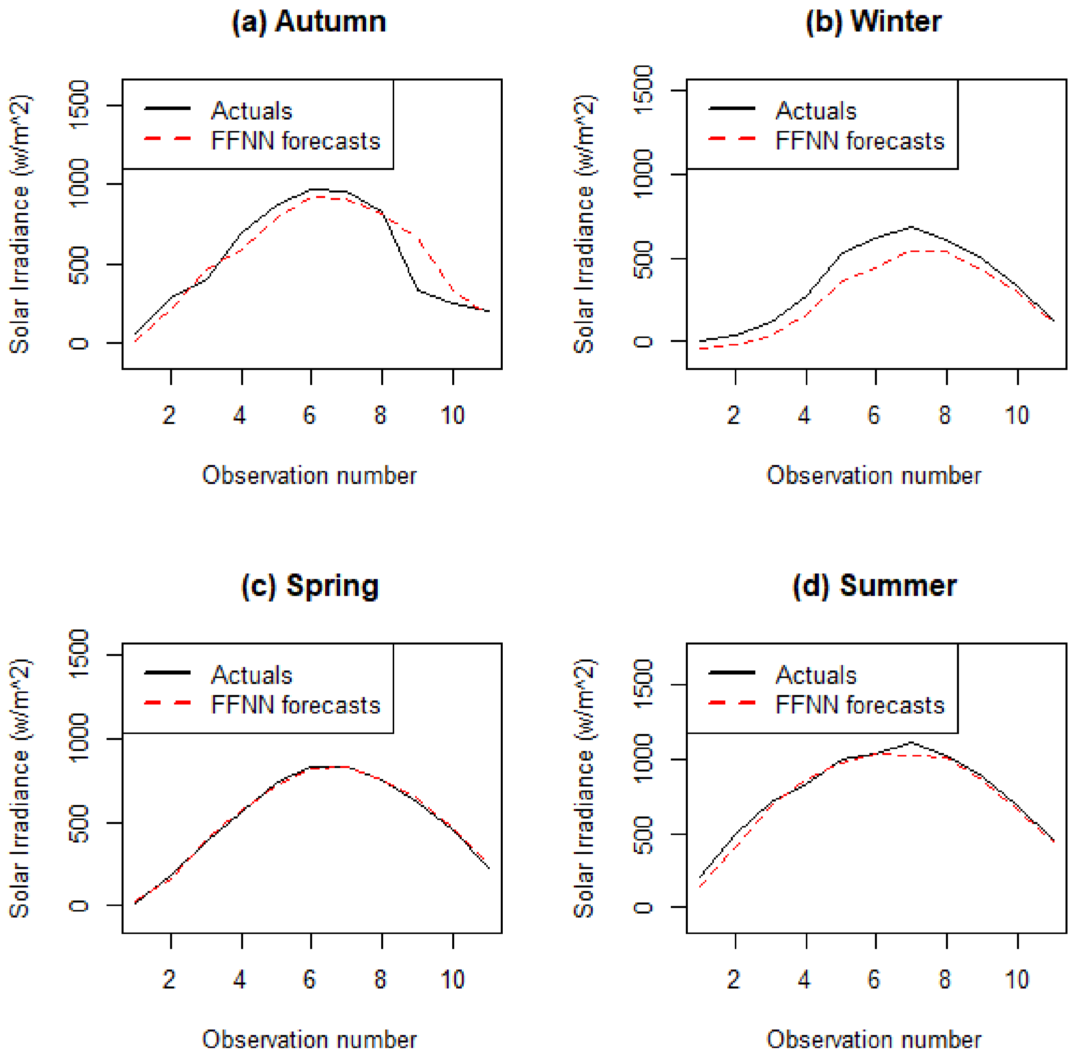


Figure 4.9: Graphical plot of forecasts (dashed line) using model M4 (FFNN) and actual hourly irradiance (solid line) for the first day of season in (a) **Top left panel**: Autumn (1 March 2018) and (b) **Top right panel**: Winter (1 June 2018) (c) **Bottom left panel**: Spring (1 September 2019) and (d) **Bottom right panel**: Summer (1 December 2019).

## 4.4 Forecast Combination

This section gives results of forecast combination of machine learning models forecasts. Two forecast combination methods are used, which are convex combination and quantile regression averaging.

### 4.4.1 Convex Combination

It is known that combining forecasts often leads to better forecast accuracy. The forecasts from different regression based time-series which are fitted above may be improved by combining them using R package called 'opera' developed by Gaillard and Goude (2016). The models developed are also referred to as experts. To combine forecasts we find aggregation of experts and analyse them by looking at the oracles. The opera package computes weights when combining the forecasts. Convex combination method works by computing the sequence of instantaneous losses suffered by the predictions from the experts (models) using loss function.

The loss function can be based on square, absolute, percentage, or pinball loss. Table 4.4 summarises the results for forecast combination models. It is giving weight 0.370 to model M2, 0.0821 to model M3, and 0.630 to model M4 when the loss function is specified as square. The pinball and absolute loss function are giving weight 0.108 to model M2, 0 to model M3, and 0.814 to model M4. From the accuracy measures Mix 2 and Mix 3 have the same least RMSE (88.732) and MAE (52.405) values compared with Mix 1.

The absolute loss suffered by the experts are given in Figure 4.10. From

Table 4.4: Evaluation of forecast combination models.

| Models (experts) | | | |
|---|---|---|---|
| | M2 | M3 | M4 |
| Square (Mix 1) | 0.108 | 0.0821 | 0.810 |
| Pinball (Mix 2) | 0.370 | 0 | 0.630 |
| Absolute (Mix 3) | 0.370 | 0 | 0.630 |
| **Accuracy measures** | | | |
| | Mix 1 | Mix 2 | Mix 3 |
| RMSE | 88.128 | 88.732 | 88.732 |
| MAE | 53.339 | 52.405 | 52.405 |

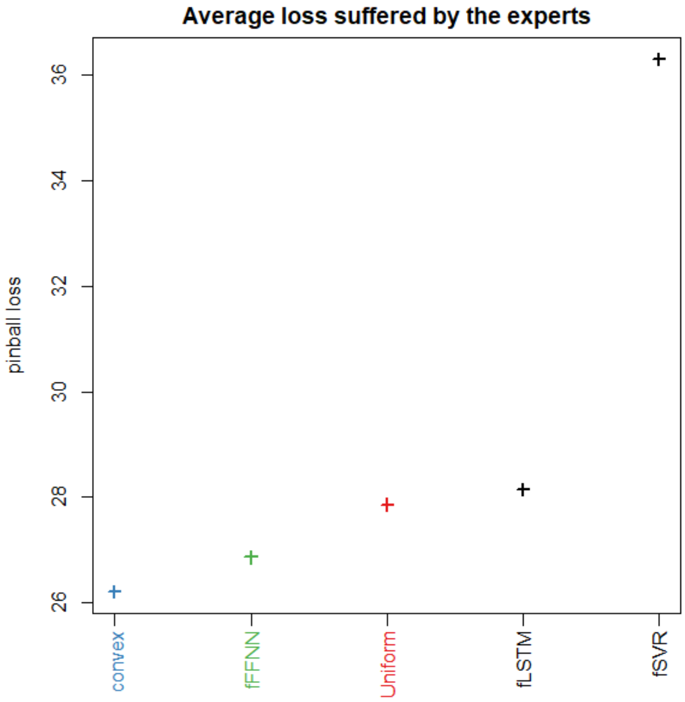


Figure 4.10: Average loss suffered by the models.

Figure 4.10 the convex combination model is shown as the best forecasting mode1 followed by model M4 (fFFNN), uniform combination (Uniform), model M2 (fLSTM), and model M3 (fSVR) respectively.

### 4.4.2 Quantile Regression Averaging

Quantile regression averaging (QRA) is another technique normally used to combine forecasts by using forecasts from each model as independent variables. The three models M2 (LSTM), M3 (SVR), and M4 (FFNN) are combined based on QRA, resulting in model M6. Model M6 (QRA) is given by:

$$y_{t,\tau}(\text{QRA}) = \beta_0 + \beta_1 \text{fM2} + \beta_2 \text{fM3} + \beta_3 \text{fM4} + \varepsilon_t \tag{4.2}$$

where fM2, fM3, and fM4 represent the forecasts from models M2, M3, and M4, respectively.

Table 4.5 gives a summary of the accuracy measures for the machine learning models, the M5 (Convex) model, and the M6 (QRA) model. Based on MAE, model M6 is the best forecasting model compared with the convex model and the machine learning models. MAE shows a slight improvement after forecast averaging. The results of absolute loss suffered by the models based on pinball losses shows M6 as the best model since it has the smallest average pinball loss value (26.017).

Combining forecasts from individual models improves the accuracy of our hourly solar irradiance predictions. Since model M6 is giving us the best forecasts we plot its forecasts in Figure 4.11. Figure 4.12 shows plots of forecasts on the first day of each season in South Africa. The forecasts are fairly close to the actual observations, especially in spring and summer.

Table 4.5: Comparative analysis of the machine learning models, Convex model, and QRA model.

| Accuracy measure | M2 | M3 | M4 | M5 (Convex) | M6 (QRA) |
|---|---|---|---|---|---|
| RMSE | 94.131 | 99.302 | 88.326 | 88.732 | 88.600 |
| MAE | 56.291 | 72.606 | 53.719 | 52.405 | 52.034 |
| Average Pinball loss | 28.146 | 36.303 | 26.859 | 26.202 | 26.017 |

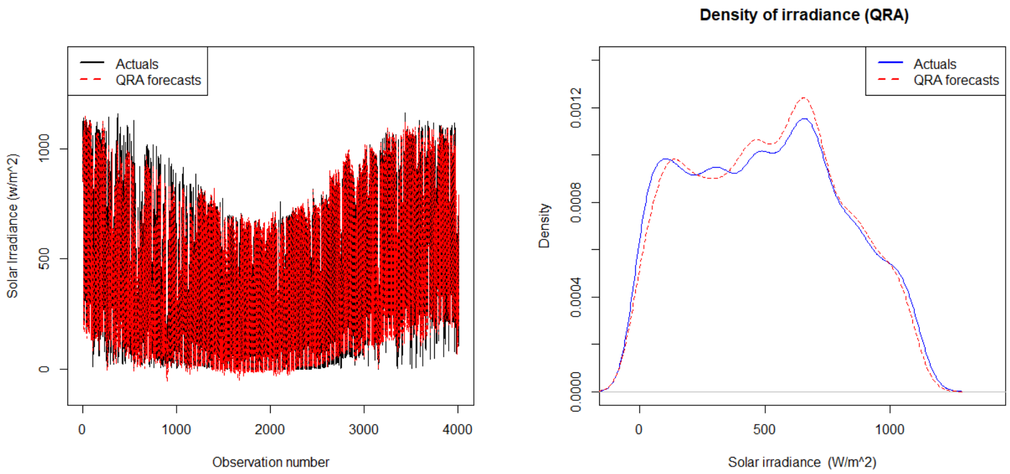


Figure 4.11: Model M6 (QRA) forecasts with density plot.

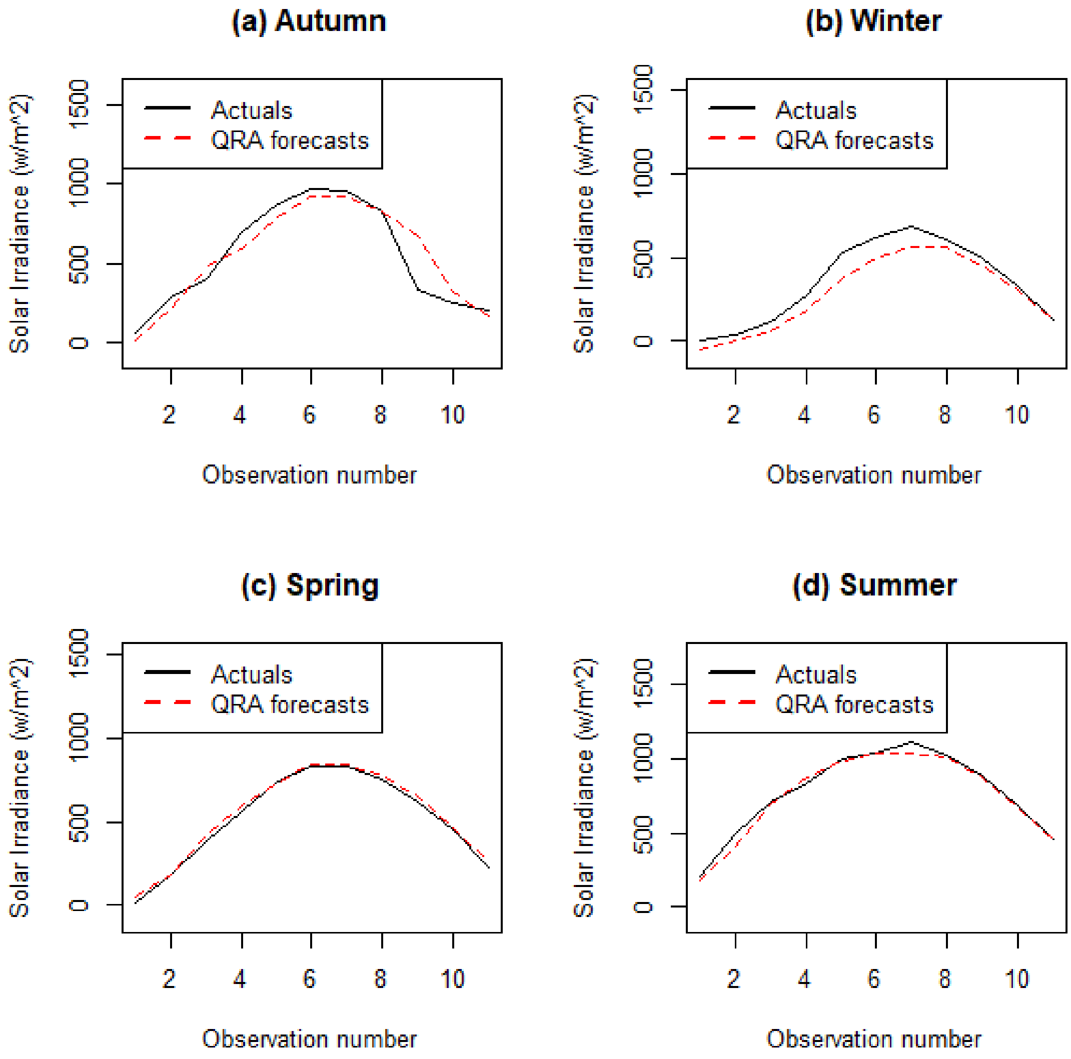


Figure 4.12: Plots of M6 (QRA) forecasts on the first day of each season in South Africa.

## 4.5 Comparative Analysis of the Models

This section presents the evaluation of the fitted models based on the empirical prediction intervals (PIs) and forecast error distributions of each model forecasts.

### 4.5.1 Evaluation of Prediction Intervals

Table 4.6 gives summary statistics of the PIWs for the models M2 to M6 for PINC value of 95% level of confidence. Model M2 has the narrowest standard deviation which indicates that it has narrower PIWs compared to models M3 to M6. All the PIWs distributions are approximately normally distributed since a normally distributed data should have a skewness value of zero. The values for kurtosis are all less than 3, showing that the distributions are all platykurtic since for a normally distributed (mesokurtic) data kurtosis value should be equal to three.

Table 4.6: Model PIWs comparisons.

| | **Mean** | **Median** | **Min** | **Max** | **St.Dev.** | **Skewness** | **Kurtosis** |
|---|---|---|---|---|---|---|---|
| M2 | 420.3 | 419.4 | 277.2 | 584.8 | 75.801 | 0.058 | -1.002 |
| M3 | 399.1 | 401.2 | 221.9 | 571.9 | 77.437 | 0.002 | -0.865 |
| M4 | 381.0 | 382.7 | 209.2 | 578.6 | 91.181 | 0.062 | -0.977 |
| M5 | 384.4 | 384.9 | 225.6 | 566.9 | 84.032 | 0.064 | -0.989 |
| M6 | 381.2 | 382.1 | 215.0 | 569.6 | 86.962 | 0.062 | -0.984 |

Figure 4.13 shows box plots of widths of the PIs for all the fitted models. From the figure, M2 has the narrowest PI compared to models M3 to model M6. Figure 4.14 density plots of widths of the PIs for the forecasting models M2 to M6. The density plots are all similar except for the PI from model

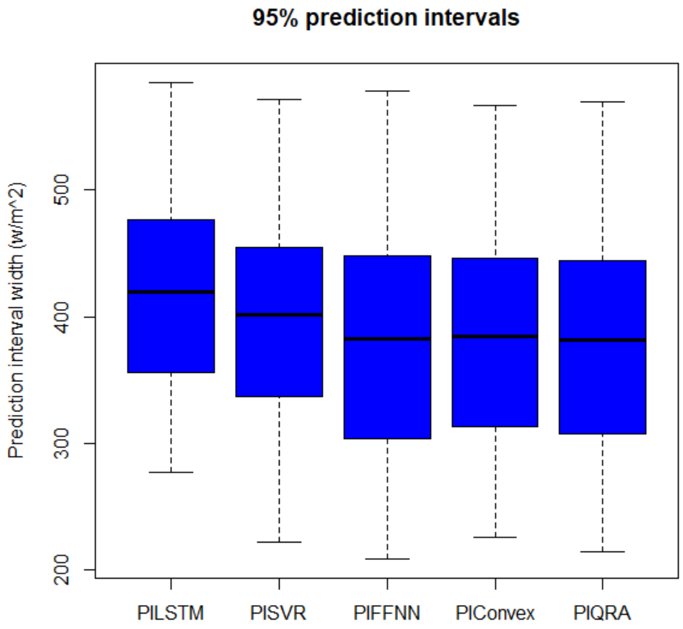


Figure 4.13: Prediction interval widths for models M2 (PILSTM), M3 (PISVR), M4 (PIFFNN), M5 (PIConvex), and M6 (PIQRA).

M3 which has the widest PI.

In order to select the best model based on the analysis of the PIWs, we need to calculate the PICPs and PINAWs including a count of the number of predictions below and above the PIs. This is done for various PINC values, which are 90%, 95% and 99%, respectively. A model with a PICP value close to the PINC value is preferred. A fitted model that has better PIWs has the lowest value of PINAW, and is one which is preferred over other fitted models. Table 4.7 shows a comparative evaluation of the models using PI indices for different PINC values. All models have valid PICPs for the three PINC

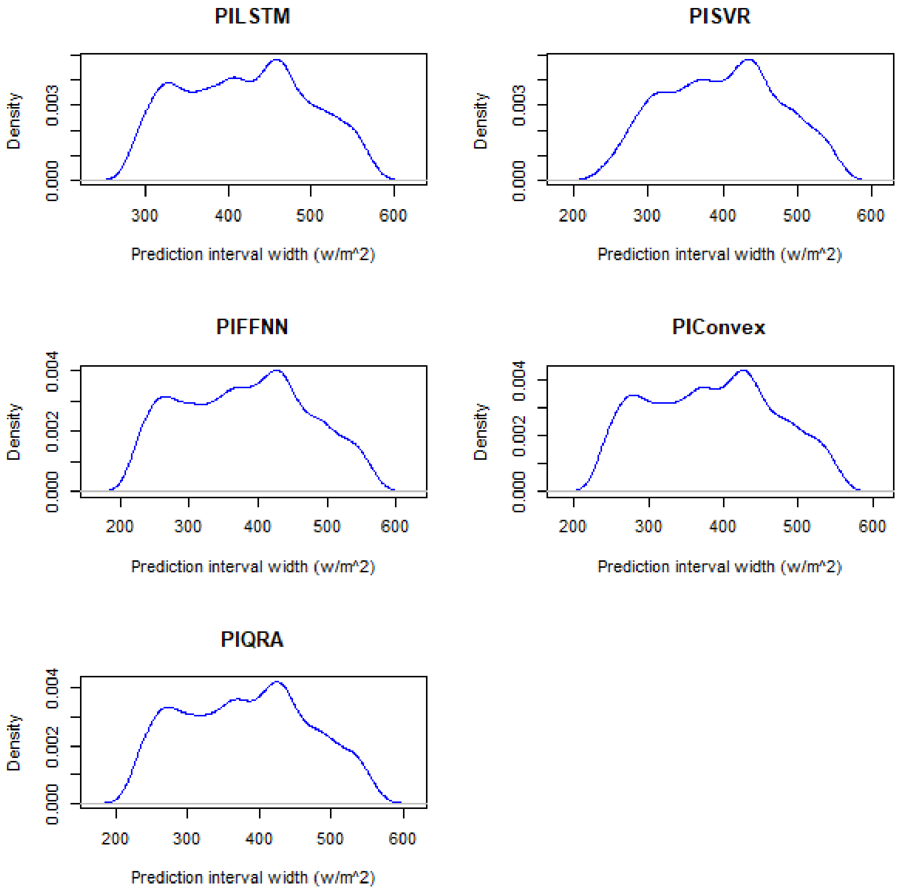


Figure 4.14: Density plots of the Prediction interval widths for models M2 (PILSTM), M3 (PISVR), M4 (PIFFNN), M5 (Convex), and M6 (PIQRA).

values. Model M2 and M3 have the same lowest PICPs at 90% and 95%, with model M3 having the lowest PICP at 99% compared to all other models. The best model based on PINAW at 90% is M6, M4 at 95%, and M3 at 99%. There is no consistency in the PICPs and PINAWs for different PINC values. However, due to results in Table 4.5, model M6 (QRA) is selected as the best model.

Table 4.7: Model PINCs comparisons.

| PINC | Model | PICP (%) | PINAW (%) | Below LL | Above UL |
|---|---|---|---|---|---|
| 90% | M2 | 90.03 | 25.58 | 200 | 200 |
| | M3 | 90.03 | 26.59 | 201 | 199 |
| | M4 | 90.05 | 23.91 | 199 | 200 |
| | M5 | 90.05 | 23.88 | 200 | 199 |
| | M6 | 90.03 | 23.67 | 200 | 200 |
| 95% | M2 | 95.04 | 36.17 | 100 | 99 |
| | M3 | 95.04 | 34.35 | 100 | 99 |
| | M4 | 95.06 | 32.79 | 99 | 99 |
| | M5 | 95.06 | 33.08 | 99 | 99 |
| | M6 | 95.06 | 32.81 | 99 | 99 |
| 99% | M2 | 99.05 | 58.03 | 19 | 19 |
| | M3 | 99.03 | 53.35 | 19 | 20 |
| | M4 | 99.05 | 54.96 | 19 | 19 |
| | M5 | 99.05 | 55.12 | 19 | 19 |
| | M6 | 99.05 | 54.94 | 19 | 19 |

## Combination of interval limits

Suppose we have $100(1-\alpha)\%$ forecast intervals $[LL_{ij}, UL_{ij}], i = 1, ..., m, j = 1, ..., k$, where $m$ is the number of forecast point and $k$ is the number of forecasters. In this study we use three basic methods for combining interval forecasts which are; average method, median method, and envelop method discussed by Gaba et al. (2017). Average (Av) limits are given by $LL_{Av} = (1/k)\sum_{j=1}^{k} LL_{ij}$ and $UL_{Av} = (1/k)\sum_{j=1}^{k} UL_{ij}$. Median (Md) limits are given by $LL_{Md} = \{LL_{i1}, ..., LL_{ik}\}$ and $UL_{Md} = \{UL_{i1}, ..., UL_{ik}\}$. Envelop (En) limits are given by $LL_{En} = \min\{LL_{i1}, ..., LL_{ik}\}$ and $UL_{En} =$

$\max\{UL_{i1}, ..., UL_{ik}\}$. Table 4.8 shows a comparative analysis of PIWs after combining the PIWs of individual models based on the average, median, and envelop methods. All combination methods have valid PICPs for the three PINC values with the median method having the lowest PINAWs for the three PINC values. Combining interval limits yield better results since the number of forecasts below and above limits decreases significantly.

Table 4.8: PINCs comparisons based on average, median, and envelop of lower and upper PIs.

| PINC | Model | PICP (%) | PINAW (%) | Below LL | Above UL |
|---|---|---|---|---|---|
| 90% | Average | 90.68 | 24.73 | 194 | 180 |
| | Median | 90.18 | 23.89 | 200 | 194 |
| | Envelop | 93.82 | 30.03 | 136 | 112 |
| 95% | Mean | 95.26 | 33.84 | 99 | 91 |
| | Median | 95.09 | 33.05 | 98 | 99 |
| | Envelop | 96.68 | 39.09 | 72 | 61 |
| 99% | Mean | 99.10 | 55.28 | 20 | 16 |
| | Median | 99.10 | 55.08 | 19 | 17 |
| | Envelop | 99.33 | 59.79 | 14 | 13 |

### 4.5.2 Residual Analysis

Summary statistics of the residuals of the models are given in Table 4.9. It can be seen that model M4 has the smallest standard deviation which indicates that it has the narrowest error distribution compared to models M3 to M6, this implies that M4 is the best compared to other models, followed by model M6. All the error distributions are skewed to the left since the values of their skewness are all negative. The values for kurtosis are all greater than 3, showing that the distributions are all leptokurtic.

Table 4.9: Model residuals comparisons.

| | Mean | Median | Min | Max | St.Dev. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|
| M2 | -6.371 | -2.965 | -711.639 | 634.962 | 93.927 | -1.006 | 8.436 |
| M3 | 19.050 | 33.090 | -637.200 | 600.870 | 97.471 | -1.263 | 5.436 |
| M4 | 3.045 | 8.223 | -662.404 | 624.515 | 88.285 | -1.181 | 8.829 |
| M5 | -0.439 | 4.169 | -680.111 | 628.380 | 88.742 | -1.158 | 9.153 |
| M6 | -5.450 | 0.000 | -681.130 | 619.630 | 88.443 | -1.213 | 9.177 |

Figure 4.15 shows box plots of the forecast errors for all the fitted models. From the figure, M4 has the narrowest error distribution compared to models M3 to M6, implying that M4 is the best model compared to other models. Figure 4.16 shows density plots of the forecast errors for the forecasting models M2 to M6. The density plots are all similar except for the forecast errors from models M2 and M3, respectively.

### 4.5.3 Percentage Improvement

Table 4.10 gives the percentage of improvement rates of the best model over other models. The percentage improvement of M6 over M2, M3, and M4 are found to be 7.56%, 28.33%, and 1.30% respectively. This means M4 (FFNN) is the second best model after M6 (QRA) and their predictive performance are very close to one another.

Table 4.10: Percentage improvement rates.

| Models | MAE (best model) | MAE (other model) | Percentage improvement |
|---|---|---|---|
| QRA & LSTM | 52.034 | 56.291 | 7.56% |
| QRA & SVR | 52.034 | 72.606 | 28.33% |
| QRA & FFNN | 52.034 | 53.719 | 1.30% |

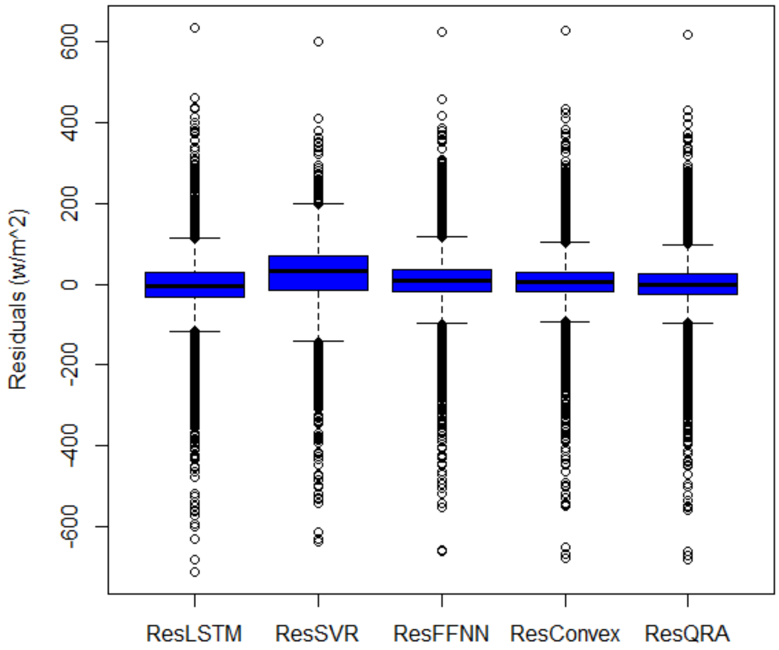


Figure 4.15: Box plots of residuals from models M2 (ResLSTM), M3 (ResSVR), M4 (ResFFNN), M5 (ResConvex), and M6 (ResQRA).

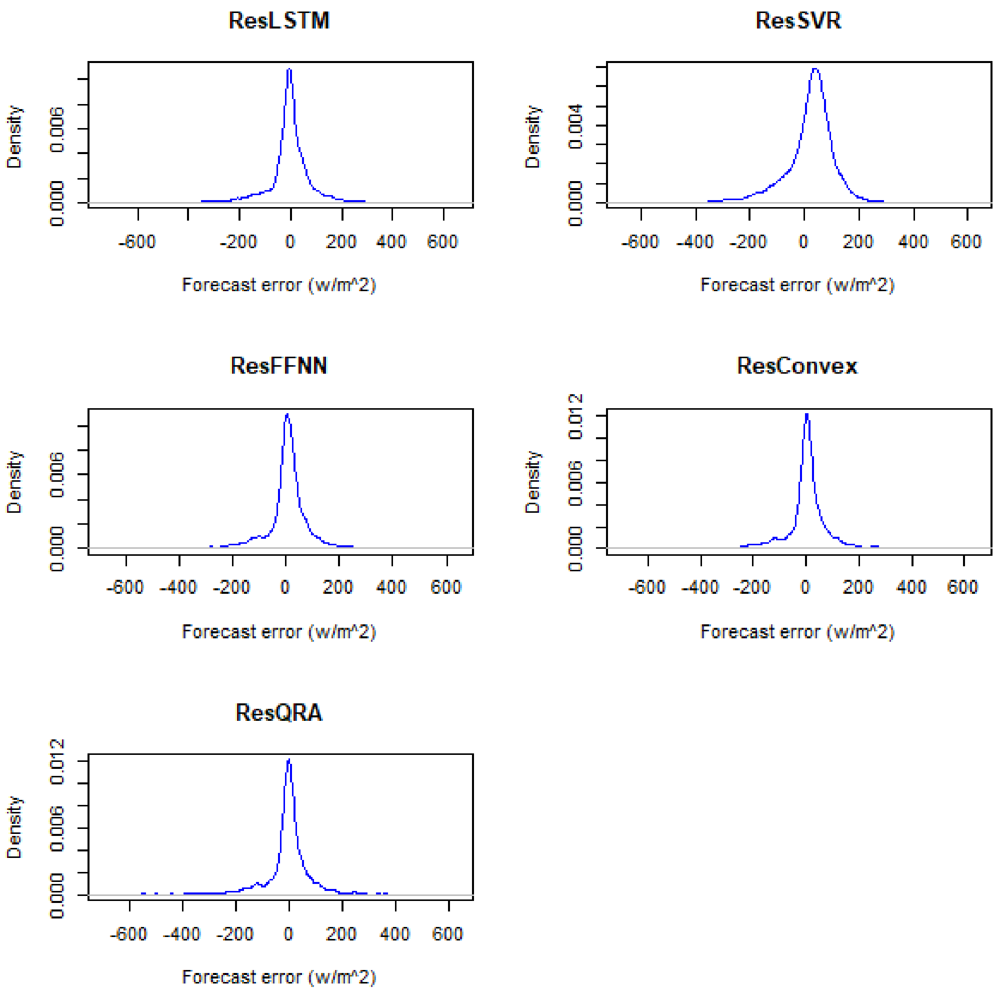


Figure 4.16: The error distribution of forecasting techniques for M2 (ResLSTM), M3 (ResSVR), M4 (ResFFNN), M5 (ResConvex), and M6 (ResQRA).

## 4.6 Discussion and Plots of out of Sample Forecasts

From the comparative analysis in Table 4.5 based on MAE and average pinball losses, M6 (QRA) is the best fitting model and can be used for predicting hourly solar irradiance. From the PIWs analysis at 95% level of confidence, M2 (LSTM) has the narrowest PI compared to models M3 to M6, implying that M2 is the best compared to other models. Further analysis of the PIWs based on the PICPs and PINAWs including a count of the number of predictions below and above the PIs shows no consistency. The best model based on PINAW at 90% is M6 (QRA), M4 (FFNN) at 95%, and M3 (SVR) at 99%.

There is no consistency in the PICPs and PINAWs for different PINC values. The residual analysis shows M4 (FFNN) as the best model with narrowest error distribution compared to other models, followed by model M6. However, due to results in Table 4.5 and Table 4.10, we stick with model M6 (QRA) as our best model. Table 4.8 shows a comparative analysis of PIWs after combining the PIWs of individual models based on the average, median, and envelop methods. All combination methods have valid PICPs for the three PINC values with the median method having the lowest PINAWs for the three PINC values. Combining interval limits yield better results since the number of forecasts below and above limits decreases significantly.

The plot of actual hourly solar irradiance superimposed with forecasted irradiance from model M6 (QRA) (1 January 2018 to 31 December 2018) given in Figure 4.17 shows that the forecasts follow hourly solar irradiance very well.

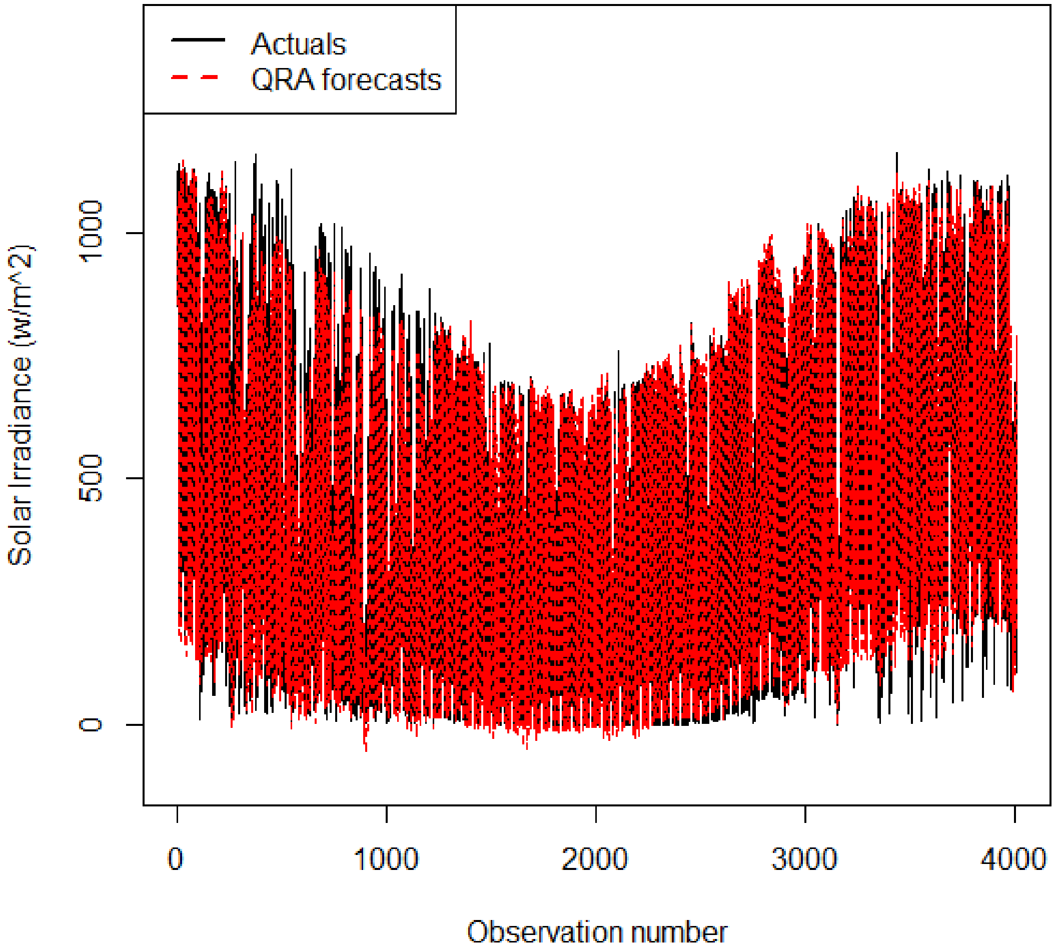


Figure 4.17: Model M6 (QRA) forecasts.

Table A.1 in Appendix A summarises the error measures for out of sample evaluation using model M6 for the month January to December 2018. The accuracy measures are significantly lower from April to August. Figures A.1-A.6 show hourly solar irradiance superimposed with forecasts together with their respective densities. The forecasts follow actual hourly solar irradiance very well, especially from June to October 2018.

# Chapter 5

# Conclusion

## 5.1 Introduction

This chapter summarises the research findings that were discussed in Chapter 4 and presents recommendations, limitations of the study and suggest areas for future research.

## 5.2 Research Findings

Solar irradiance forecasting is important for programming back-up, operational planning, short-term power purchases, switching other energy sources, planning for reserve usage and peak load demand. This research project focused on forecasting hourly solar irradiance (global horizontal irradiance (GHI)) using the irradiance data for the period 2014 to 2018 obtained from the Southern African Universities Radiometric Network (SAURAN) database for University of Pretoria radiometric (UPR) station.

Modelling hourly solar irradiance using long short-term memory (LSTM) networks, support vector regression (SVR), and feed forward neural networks (FFNN) models were discussed in Chapter 4. Before modelling hourly solar irradiance, the variable selection was done using the least absolute shrinkage and selection operator (Lasso) which showed that wind direction and day of the month are not important in forecasting solar irradiance amongst other features, and were therefore not considered in the modelling stage.

According to findings in Chapter 4, among the three fitted machine learning models in Section 4.3, the FFNN model produced the best forecast accuracy based on the MAE and RMSE. Later, the forecasts from the machine learning models were combined in Section 4.4 using the convex combination method and quantile regression averaging (QRA). Based on MAE and average pinball losses, the QRA model was found to be the best forecast combination method, and also the best forecasting model compared with the machine learning models.

From the prediction interval widths (PIWs) analysis in Section 4.5, at 95% level of confidence M2 (LSTM) has the narrowest PI compared to model M3 to model M6, this implies that M2 is the best compared to other models. Further analysis of the PIWs based on the prediction interval coverage probabilities (PICPs), and prediction interval normalised average widths (PINAWs) including a count of the number of predictions below and above the PIs showed inconsistency. The best model based on PINAW at 90% is M6 (QRA), M4 (FFNN) at 95%, and M3 (SVR) at 99%. There is inconsistency in the results of PICPs and PINAWs for different PINC values. In

Section 4.5.1, a comparative analysis of PIWs after combining the PIWs of individual models based on the average, median, and envelop methods was done. The results show the all combination methods have valid PICPs for the three PINC values with the median method having the lowest PINAWs for the three PINC values. Combining interval limits yield better results since the number of forecasts below and above limits decreases significantly. The residual analysis in Section 4.5.2 shows M4 (FFNN) as the best model with narrowest error distribution compared to other models.

Finally, in Section 4.5.3, the percentage improvement rates of the best model over other models were calculated. The percentage improvement of M6 over M2, M3, and M4 were found to be 7.56%, 28.33%, and 1.30% respectively. This means M4 (FFNN) is the second best model after M6 (QRA) and their predictive performance is very close to one another. However, based on the results from MAE and average pinball losses, model M6 (QRA) was selected as the as the best forecasting model. The median method for combining interval limits gave the best results on PIWs analysis.

## 5.3 Recommendations

The information derived from this research study on forecasting solar irradiance is important to electricity utility decision-makers in South Africa in balancing demand and supply of electricity in an effective way which favours future economic prosperity and environmental security. This research study recommends forecast combination and combining interval limits for forecasting solar irradiance in South Africa.

## 5.4 Limitations of the Study

One of the limitations of this research study is that only some of the weather variables were used as covariates to model solar irradiance. It would have been interesting to forecast hourly solar irradiance with the inclusion of other predictor variables such as precipitation and cloud cover. Another limitation of this study is that only one radiometric station in South Africa out of 19 radiometric stations which are owned by SAURAN was studied.

## 5.5 Future Research

Future work will focus on including more radiometric stations based in South Africa and also consider different weather conditions at a coastal and inland region by doing spatial analysis.

# Appendix A

# Forecasts for the Months January to December 2018

Table A.1: Forecast accuracy measures root mean square error (RMSE), mean absolute error (MAE) for the forecasts of January to December 2018.

| **Month** | **RMSE** | **MAE** |
|---|---|---|
| January | 103.389 | 69.195 |
| February | 124.483 | 91.413 |
| March | 103.091 | 75.659 |
| April | 92.043 | 58.864 |
| May | 59.154 | 33.180 |
| June | 34.847 | 20.039 |
| July | 44.066 | 24.256 |
| August | 40.814 | 23.729 |
| September | 53.305 | 30.064 |
| October | 98.832 | 51.750 |
| November | 118.346 | 71.924 |
| December | 121.898 | 77.909 |

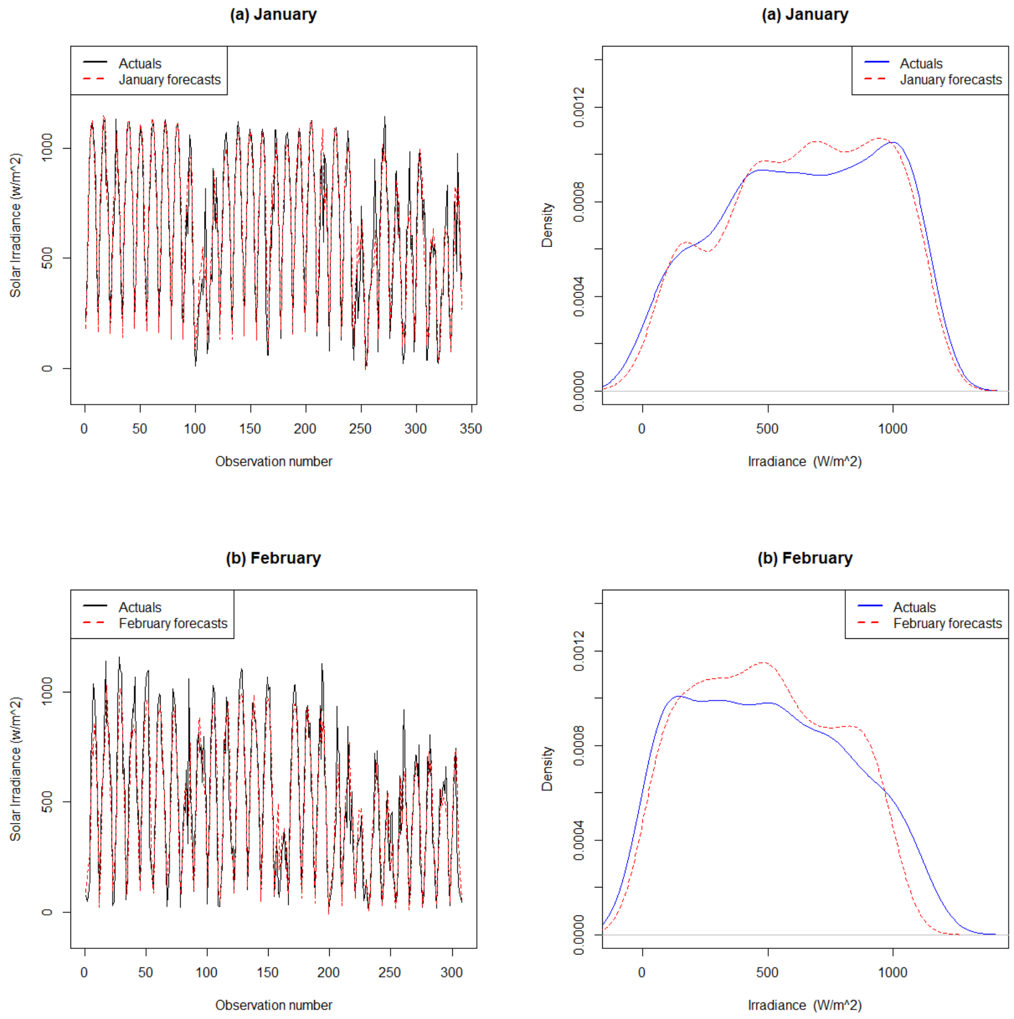


Figure A.1: Hourly irradiance with forecasts for the months January and February 2018.

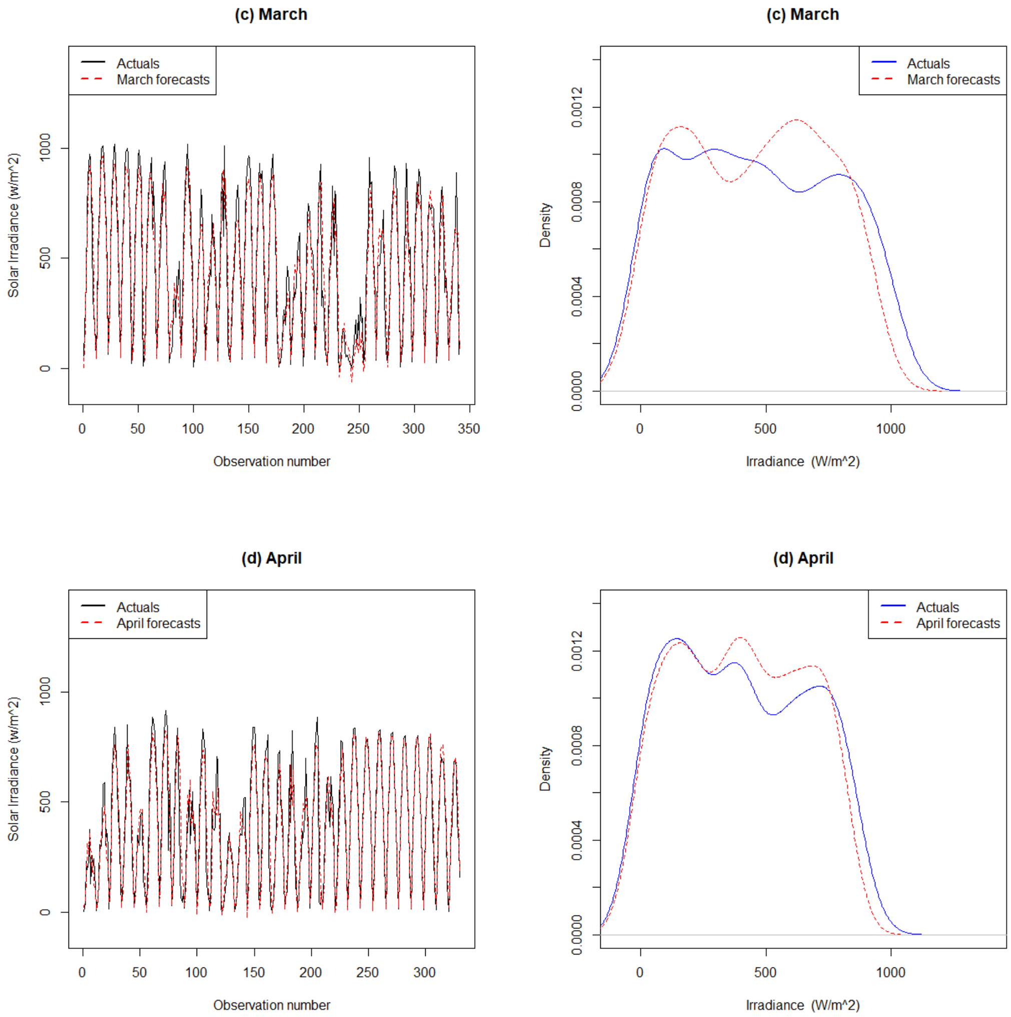


Figure A.2: Hourly irradiance with forecasts for the months March and April 2018.

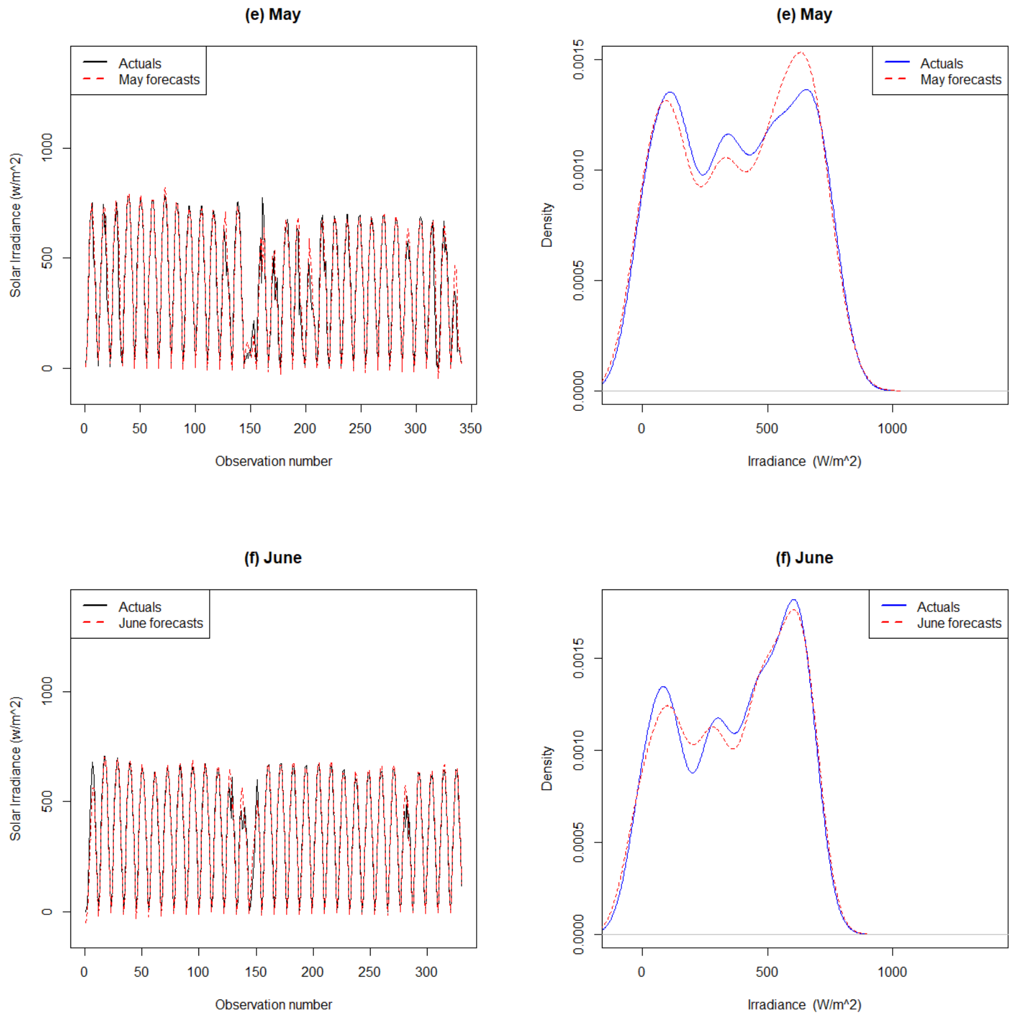


Figure A.3: Hourly irradiance with forecasts for the months May and June 2018.

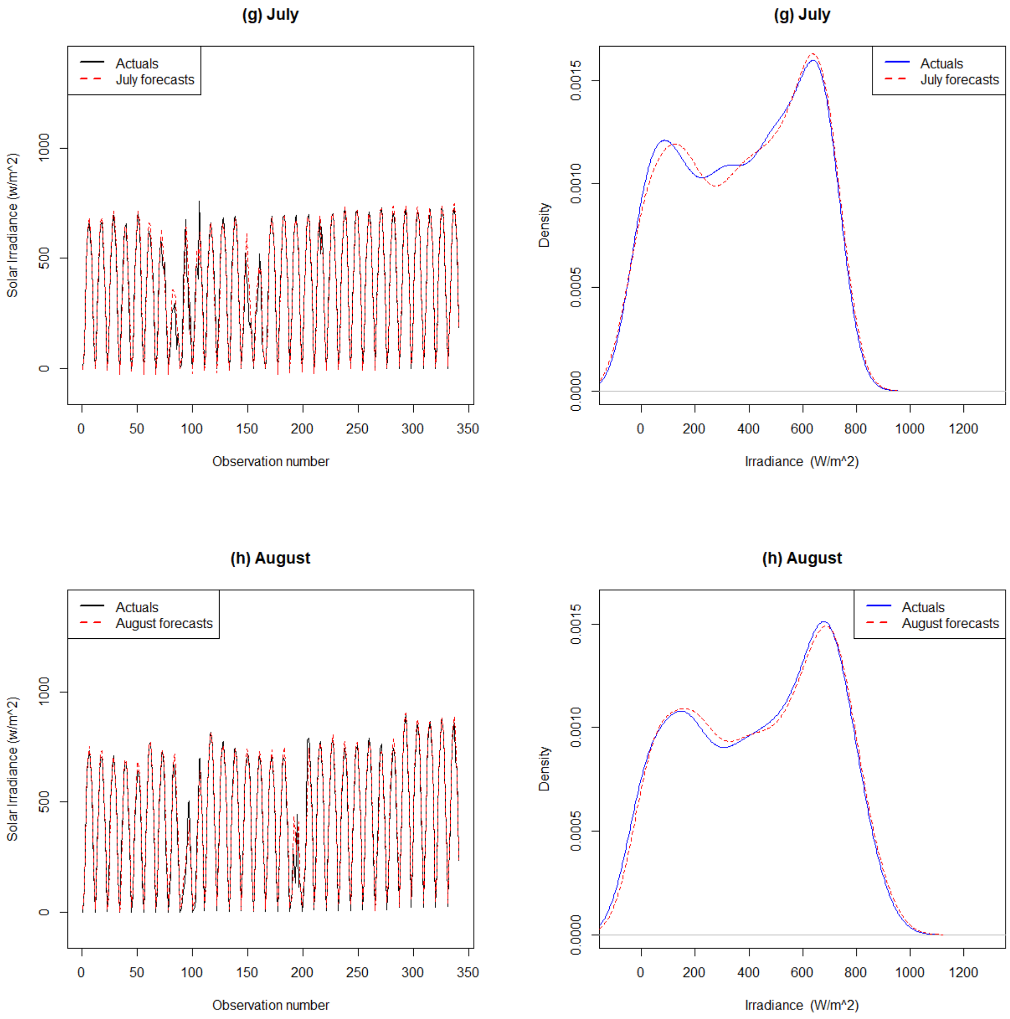


Figure A.4: Hourly irradiance with forecasts for the months July and August 2018.

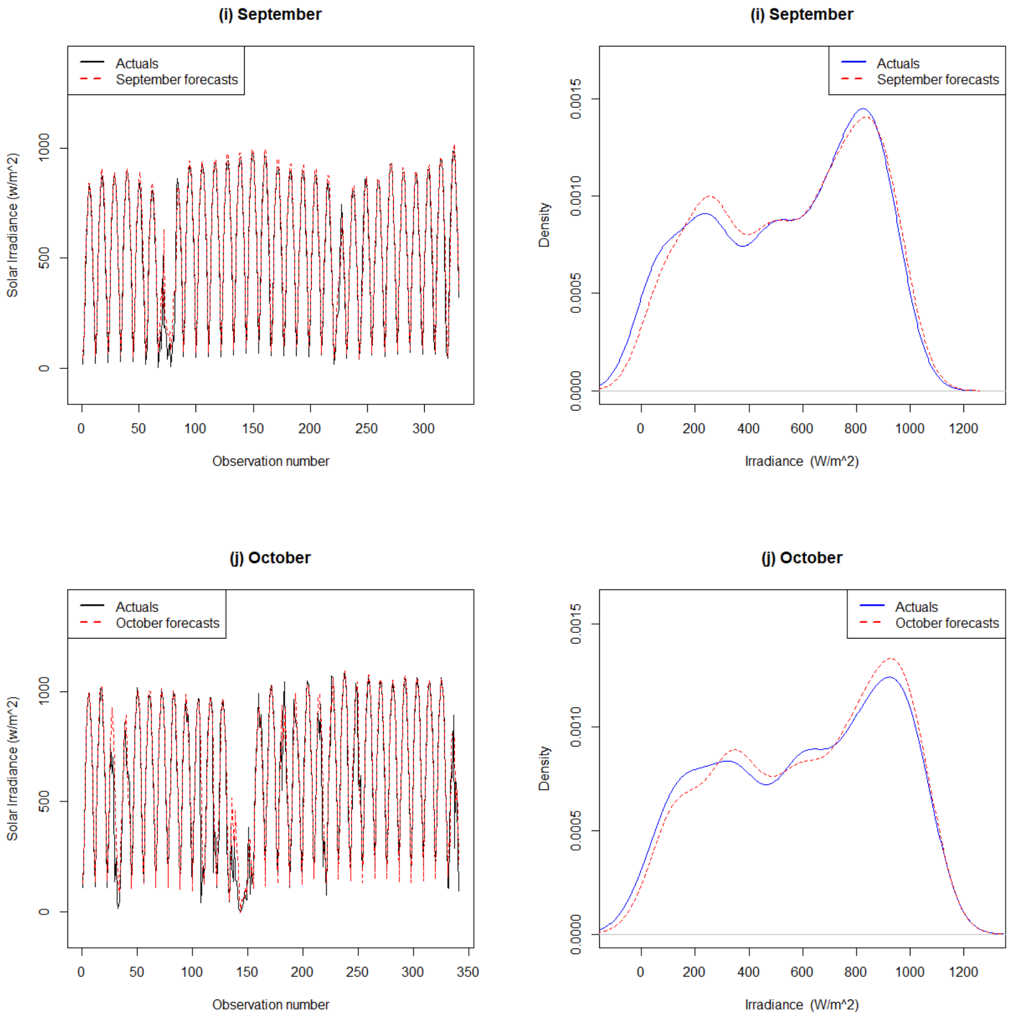


Figure A.5: Hourly irradiance with forecasts for the months September and October 2018.

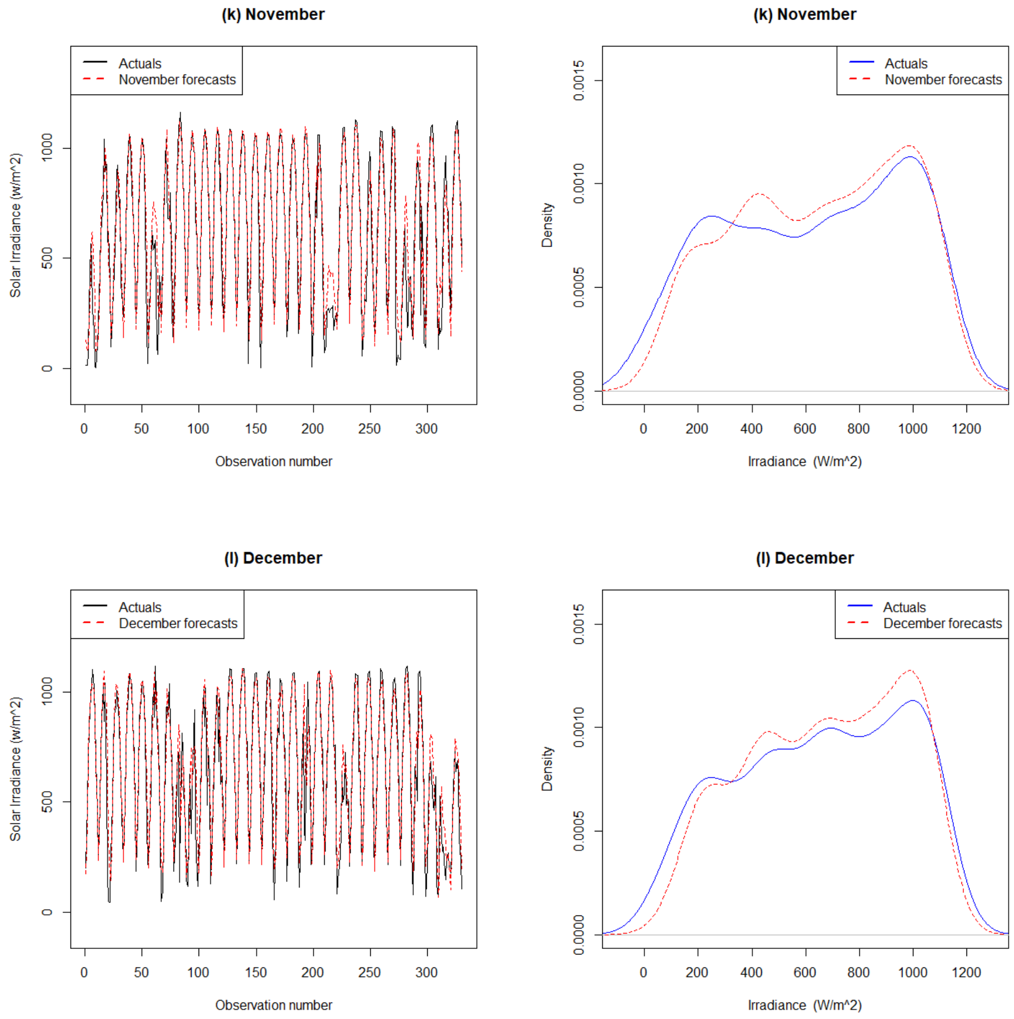


Figure A.6: Hourly irradiance with forecasts for the months Noveber and December 2018.

# Appendix B

# Some selected R Code

```
# The following packages in R are used:
# forecast, ggplot2, qgam, mgcv, tseries, e1071, glmnet, hierNet.
######################################################################
library(forecast)
library(ggplot2)
library(tseries)
library(e1071)
library(glmnet)
library(hierNet)
#####################################################################
# time series, qqnorn, density and box plot for solar irradiance
#####################################################################
attach(UPRanlyticdata)
head(UPRanlyticdata)
win.graph()
par(mfrow=c(2,2))
A<-ts(GHI)
plot(A,xlab="Observation number",ylab="Solar irradiance (W/m^2)"
,main="(a) Plot of irradiance",col = "blue")
plot(density(A),xlab="Solar irradiance (W/m^2)
",main="(b) Density plot",col = "blue")
qqnorm(A,col = "blue",main="(c) Normal QQ plot")
qqline(A)
boxplot(A,main="(d) Box plot",varwidth=TRUE,
xlab="Solar irradiance (W/m^2)", col = "blue",horizontal= TRUE)
###################################################################
## calculating summary statistics, skewness and kurtosis of GHI
####################################################################
summary(A)
library(e1071)
sd(A)
skewness(A)
kurtosis(A)
###################################################
## Fitting and extracting Non-linear Trend values
#####################################################
z <- ts(GHI)
```

```
win.graph()
plot(z, xlab="Observation number",ylim=c(0,1200),type="l",
ylab="Solar irradiance (W/m^2)")
length(GHI)
r=smooth.spline(time(z), z)
r  # 0.1112481
lines(smooth.spline(time(z), z, spar= 0.1112481), lwd=3,col="red")
dpdfits = fitted((smooth.spline(time(z), z, spar= 0.1112481)))
write.table(dpdfits,"~/GHIfittedspline.txt",sep="\t")
##########################################################
## PRINCIPAL COMPONENT REGRESSION (PCR)
##########################################################
attach(UPRnewdata) # Data will all features
head(UPRnewdata)
library(pls)

GHI_data_test <- 16045:nrow(UPRnewdata)
GHI_data_test
data_train <- UPRnewdata[-GHI_data_test, ]
data_train
data_test <- UPRnewdata[GHI_data_test, ]
data_test

m1 <- pcr(GHI~Temp+BP+RH+RainTot+WS+WD+WD_SD+Lag1+Lag2+noltrend+Hour+Month+Day,
ncomp = 12, data = data_train, validation = "CV")
summary(m1)
forecasts <- predict(m1, ncomp = 1:12, newdata = data_test)
forecasts <- data.frame(forecasts)
forecasts
write.table(forecasts,"~/forecastsPCA.txt",sep="\t")
f12 <- forecasts$GHI.12.comps
x <- data_test$GHI
accuracy(f12,x)

#######################################
## FORECAST COMBINATION OPERA
######################################
library(forecast)
attach(ForecastsUPRnew) # Forecasts from all ML models
head(ForecastsUPRnew)
win.graph()
accuracy(Flstm, GHI)
accuracy(Fsvr, GHI)
accuracy(Fffnn, GHI)
########################################
Y <-GHI
X <- cbind(Flstm, Fsvr, Fffnn)
matplot(cbind(Y,X), type="l", col=1:4)
fLSTM<- Flstm
fSVR<- Fsvr
fFFNN<- Fffnn
X <- cbind(fLSTM, fSVR, fFFNN)
# How good are the expert? Look at the oracles

library(opera)
oracle.convex <- oracle(Y = Y, experts = X, loss.type = "pinball",
model = "convex")
oracle.convex
plot(oracle.convex)
print(oracle.convex) # print is same as plot in results

mix1 =  0.37*Flstm + 4.99e-21*Fsvr + 0.63*Fffnn #  pinball
mix2 = 0.108*Flstm + 0.0821*Fsvr +0.81*Fffnn # square
mix3= 0.37*Flstm + 4.86e-21*Fsvr+ 0.63*Fffnn #  absolute
# Accuracy measures
accuracy(mix1, GHI)
accuracy(mix2, GHI)
accuracy(mix3, GHI)

mix1
write.table(mix1,"~/Fconvex.txt",sep="\t")
#################################################
## QUANTILE REGRESSION AVERAGING
#################################################
attach(ForecastsUPRnew)
head(ForecastsUPRnew)
win.graph()
y <- ts(GHI)
plot(y, xlab="Observation number", ylab="Hourly irradiance")

library(quantreg)
qr.GHI = rq(GHI ~ Flstm+Fsvr+Fffnn,data= ForecastsUPRnew, tau=0.5) #tau = 0.025, 0.5, 0.975

summary.rq(qr.GHI,se="boot") # can use se = "nid" or se="ker"
lines(qr.GHI$fit, col="red")
fQRA = fitted(qr.GHI)
write.table(fQRA,"~/QRA05.txt",sep="\t") #LL0025, QRA05, UL0975
accuracy(fQRA,GHI)
###########################################################
## Pinball loss function using r-package gefcom2017
###########################################################
#install.packages("remotes")
#remotes::install_github("camroach87/gefcom2017")
attach(ForecastsUPRnew) ## Load all forecasts
library(gefcom2017)
pinball_loss <- function(tau, y, q) {
pl_df <- data.frame(tau = tau,
y = y,
q = q)

pl_df <- pl_df %>%
mutate(L = ifelse(y>=q,
tau/100 * (y-q),
(1-tau/100) * (q-y)))

return(pl_df)
}

tau= 50
y= GHI
q= Fconvex   # Flstm, Fsvr, Fffnn, Fconvex, Fqra
z = pinball_loss(tau, y, q)
```

```
z
write.table(z,"~/pinballFconvex.txt",sep="\t")
qloss =z$L
a=ts(qloss)
plot(a)
mean(qloss)

### Forecast plot with 2 lines
### Forecast plot for LSTM, SVR, FFNN, Convex, QRA, PCR
## Flstm, Fsvr, Fffnn, Fconvex, Fqra, f12
y=ts(GHI)
win.graph()
plot(y,xlab="Observation number",ylim=c(-100,1400),type="l",
ylab="Solar Irradiance (w/m^2)")
lines(Fsvr,col="red", lty=2)
legend("topleft",col=c("black","red"), lty=1:2,lwd=2,
legend=c("Actuals", "SVR forecasts"))

# Density plot with 2 lines
### Density plot for LSTM, SVR, FFNN, Convex, QRA, PCR
## Flstm, Fsvr, Fffnn, Fconvex, Fqra, f12
x1= density(GHI)
plot(x1,ylim=c(0.0,0.0014), xlim=c(-100,1400),col="blue", main=" Density of irradiance (LSTM)",
xlab="Solar irradiance  (W/m^2)")
x2=density(Flstm)
lines(x2,col="red", lty=2)
legend("topright",lty=c(1,2), lwd=2,col=c("blue","red"),
legend=c("Actuals", "LSTM forecasts"))
####################################################################
### Prediction Interval Width for all models
#####################################################################

## Flstm, Fsvr, Fffnn, Fconvex, Fqra
qr.GHI = rq(GHI ~ Flstm,data= ForecastsUPRnew, tau=0.05) #LL = 0.05, 0.025, 0.005
#UP = 0.950, 0.975, 0.995
summary.rq(qr.GHI,se="boot") # can use se = "nid" or se="ker"
#lines(qr.GHI$fit, col="red")

fQRA = fitted(qr.GHI)
write.table(fQRA,"~/Flstm005.txt",sep="\t") #LL0025, QRA05, UL0975
accuracy(fQRA,GHI)

######################################################################
#### Model PIWs comparisons at 95%
######################################################################
attach(PIs_UPRnew)
head(PIs_UPRnew)

PIW = c("PILSTM","PISVR","PIFFNN","PIConvex","PIQRA")
win.graph()
boxplot(PIlstm95, PIsvr95, PIffnn95, PIconvex95,PIqra95, names= PIW, horizontal = FALSE,main="95% prediction intervals",
ylab="Prediction interval width (w/m^2)", col = "blue")

win.graph()
```

```r
par(mfrow=c(3,2))
plot(density(PIlstm95),xlab="Prediction interval width (w/m^2)", col="blue", main="PILSTM")
plot(density(PIsvr95),xlab="Prediction interval width (w/m^2)", col="blue", main="PISVR")
plot(density(PIffnn95),xlab="Prediction interval width (w/m^2)", col="blue", main="PIFFNN")
plot(density(PIconvex95),xlab="Prediction interval width (w/m^2)", col="blue",
main="PIConvex")
plot(density(PIqra95),xlab="Prediction interval width (w/m^2)", col="blue", main="PIQRA")

## Summary statistics for PIWs
## PIlstm95, PIsvr95, PIffnn95, PIconvex95, PIqra95
library(e1071)
summary(PIqra95)
sd(PIqra95)
skewness(PIqra95)
kurtosis(PIqra95)

#################################################################
#Residual Error Analysis
#################################################################
attach(ForecastsUPRnew)
head(ForecastsUPRnew)
ResLSTM= ForecastsUPRnew$GHI - ForecastsUPRnew$Flstm
ResSVR= ForecastsUPRnew$GHI - ForecastsUPRnew$Fsvr
ResFFNN= ForecastsUPRnew$GHI - ForecastsUPRnew$Fffnn
ResConvex= ForecastsUPRnew$GHI - ForecastsUPRnew$Fconvex
ResQRA= ForecastsUPRnew$GHI - ForecastsUPRnew$Fqra

## Summary Statistics for forecast errors
## ResLSTM, ResSVR, ResFFNN, ResConvex, ResQRA
library(e1071)
summary(ResQRA)
sd(ResQRA)
skewness(ResQRA)
kurtosis(ResQRA)

#Residual box-pot and density plot

RESID = c("ResLSTM","ResSVR","ResFFNN","ResConvex", "ResQRA")
win.graph()
boxplot(ResLSTM, ResSVR, ResFFNN, ResConvex,ResQRA, names= RESID, horizontal = FALSE,main="",
ylab="Residuals (w/m^2)", col = "blue")

win.graph()
par(mfrow=c(3,2))
plot(density(ResLSTM),xlab="Forecast error (w/m^2)", col="blue", main="ResLSTM")
plot(density(ResSVR),xlab="Forecast error (w/m^2)", col="blue", main="ResSVR")
plot(density(ResFFNN),xlab="Forecast error (w/m^2)", col="blue", main="ResFFNN")
plot(density(ResConvex),xlab="Forecast error (w/m^2)", col="blue", main="
    ResConvex")
plot(density(ResQRA),xlab="Forecast error (w/m^2)", col="blue", main="ResQRA
    ")
```

# References

Adeala, A. A., Huan, Z. and Enweremadu, C. C. (2015), ‘Evaluation of global solar radiation using multiple weather parameters as predictors for south africa provinces’, *Thermal Science* **19**(suppl. 2), 495–509.

Akarslan, E. and Hocaoglu, F. O. (2016), ‘A novel adaptive approach for hourly solar radiation forecasting’, *Renewable Energy* **87**, 628–633.

Akarslan, E., Hocaoglu, F. O. and Edizkan, R. (2018), ‘Novel short term solar irradiance forecasting models’, *Renewable Energy* **123**, 58–66.

Badescu, V. (2014), *Modeling solar radiation at the earth’s surface*, Springer.

Benmouiza, K. and Cheknane, A. (2013), ‘Forecasting hourly global solar radiation using hybrid k-means and nonlinear autoregressive neural network models’, *Energy Conversion and Management* **75**, 561–569.

Bien, J., Taylor, J. and Tibshirani, R. (2013), ‘A lasso for hierarchical interactions’, *Annals of Statistics* **41**(3), 1111.

Cristaldi, L., Leone, G. and Ottoboni, R. (2017), A hybrid approach for solar radiation and photovoltaic power short-term forecast, *in* ‘Instrumentation and Measurement Technology Conference (I2MTC), 2017 IEEE International’, IEEE, pp. 1–6.

Deb, C., Lee, S. E. and Santamouris, M. (2018), ‘Using artificial neural networks to assess hvac related energy saving in retrofitted office buildings’, *Solar Energy* **163**, 32–44.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A. J. and Vapnik, V. (1997), Support vector regression machines, *in* ‘Advances in Neural Information Processing Systems’, pp. 155–161.

Fan, J., Wang, X., Wu, L., Zhou, H., Zhang, F., Yu, X., Lu, X. and Xiang, Y. (2018), ‘Comparison of support vector machine and extreme gradient boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in china’, *Energy Conversion and Management* **164**, 102–111.

Gaba, A., Tsetlin, I. and Winkler, R. L. (2017), ‘Combining interval forecasts’, *Decision Analysis* **14**(1), 1–20.

Gaillard, P. and Goude, Y. (2015), Forecasting electricity consumption by aggregating experts; how to design a good set of experts, *in* ‘Modeling and stochastic learning for forecasting in high dimensions’, Springer, pp. 95–115.

Gaillard, P. and Goude, Y. (2016), *opera: Online Prediction by Expert Aggregation*. R package version 3.1.0.
**URL:** *https://CRAN.R-project.org/package=opera*

Gensler, A., Henze, J., Sick, B. and Raabe, N. (2016), Deep learning for solar power forecastingan approach using autoencoder and lstm neural networks, *in* ‘Systems, Man, and Cybernetics (SMC), 2016 IEEE International Conference on’, IEEE, pp. 002858–002865.

Halabi, L. M., Mekhilef, S. and Hossain, M. (2018), ‘Performance evaluation of hybrid adaptive neuro-fuzzy inference system models for predicting monthly global solar radiation’, *Applied Energy* **213**, 247–261.

Haykin, S. S., Haykin, S. S., Haykin, S. S. and Haykin, S. S. (2009), *Neural networks and learning machines*, Vol. 3, Pearson Upper Saddle River, NJ, USA:.

Hochreiter, S. and Schmidhuber, J. (1997), ‘Long short-term memory’, *Neural computation* **9**(8), 1735–1780.

Huang, J., Korolkiewicz, M., Agrawal, M. and Boland, J. (2013), ‘Forecasting solar radiation on an hourly time scale using a coupled autoregressive and dynamical system (cards) model’, *Solar Energy* **87**, 136–149.

Khatib, T. and Elmenreich, W. (2015), ‘A model for hourly solar radiation data generation from daily solar radiation data using a generalized regression artificial neural network’, *International Journal of Photoenergy* **2015**.

Li, F.-F., Wang, S.-Y. and Wei, J.-H. (2018), ‘Long term rolling prediction model for solar radiation combining empirical mode decomposition (emd) and artificial neural network (ann) techniques’, *Journal of Renewable and Sustainable Energy* **10**(1), 013704.

Liu, B. Y. and Jordan, R. C. (1960), ‘The interrelationship and characteristic distribution of direct, diffuse and total solar radiation’, *Solar Energy* **4**(3), 1–19.

Liu, R., Kuang, J., Gong, Q. and Hou, X. (2003), ‘Principal component regression analysis with spss’, *Computer Methods and Programs in Biomedicine* **71**(2), 141–147.

Maciejowska, K., Nowotarski, J. and Weron, R. (2016), 'Probabilistic forecasting of electricity spot prices using factor quantile regression averaging', *International Journal of Forecasting* **32**(3), 957–965.

McCulloch, W. S. and Pitts, W. (1943), 'A logical calculus of the ideas immanent in nervous activity', *The Bulletin of Mathematical Biophysics* **5**(4), 115–133.

Minsky, M. and Papert, S. (1990), 'Perceptrons. 1969', *Cited on* p. 1.

Mohammadi, K., Shamshirband, S., Danesh, A. S., Abdullah, M. S. and Zamani, M. (2016), 'Temperature-based estimation of global solar radiation using soft computing methodologies', *Theoretical and Applied Climatology* **125**(1-2), 101–112.

Mpfumali, P., Sigauke, C., Bere, A. and Mulaudzi, S. (2019), 'Day ahead hourly global horizontal irradiance forecastingapplication to south african data', *Energies* **12**(18), 3569.

Paoli, C., Voyant, C., Muselli, M. and Nivet, M.-L. (2010), 'Forecasting of preprocessed daily solar radiation time series using neural networks', *Solar Energy* **84**(12), 2146–2160.

Rezrazi, A., Hanini, S. and Laidi, M. (2016), 'An optimisation methodology of artificial neural network models for predicting solar radiation: a case study', *Theoretical and Applied Climatology* **123**(3-4), 769–783.

Rosenblatt, F. (1958), 'The perceptron: a probabilistic model for information storage and organization in the brain.', *Psychological Review* **65**(6), 386.

Shamshirband, S., Mohammadi, K., Tong, C. W., Zamani, M., Motamedi, S. and Ch, S. (2016), 'A hybrid svm-ffa method for prediction of monthly mean global solar radiation', *Theoretical and Applied Climatology* **125**(1-2), 53–65.

Sreekumar, S., Sharma, K. C. and Bhakar, R. (2016), Optimized support vector regression models for short term solar radiation forecasting in smart environment, *in* 'Region 10 Conference (TENCON), 2016 IEEE', IEEE, pp. 1929–1932.

Sun, S., Wang, S., Zhang, G. and Zheng, J. (2018), 'A decomposition-clustering-ensemble learning approach for solar radiation forecasting', *Solar Energy* **163**, 189–199.

Sun, X., Wang, Z. and Hu, J. (2017), 'Prediction interval construction for byproduct gas flow forecasting using optimized twin extreme learning machine', *Mathematical Problems in Engineering* **2017**.

Yang, X., Jiang, F. and Liu, H. (2013), 'Short-term solar radiation prediction based on svm with similar data, in the proceeding of 2013 ieee rpg (2nd iet renewable power generation conference)'.

Zendehboudi, A., Baseer, M. and Saidur, R. (2018), 'Application of support vector machine models for forecasting solar and wind energy resources: A review', *Journal of Cleaner Production* **199**, 272–285.

Zhandire, E. (2017), 'Solar resource classification in south africa using a new index', *Journal of Energy in Southern Africa* **28**(2), 61–70.